

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 29, 2004, 00:26:49 ; Search time 4909 Seconds
(without alignments)
11045.454 Million cell updates/sec

Title: US-10-603-141-1
Perfect score: 1251
Sequence: 1 gaatcggcagcaggcaga.....gtatcatggagagcctgac 1251

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.ats.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.ats.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sv.*
- 39: em.htgo.hum.*
- 40: em.htgo.mus.*
- 41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1251	100.0	1251	10	AF301018	AF301018 Mus muscu
2	1229.8	98.3	1823	10	AF305709	AF305709 Mus muscu
3	1053	84.2	1053	6	E28986	E28986 Seven-pass
4	1053	84.2	1053	6	E28987	E28987 Seven-pass
5	1009.4	80.7	231959	2	AC126907	AC126907 Rattus no
6	1009.4	80.7	254420	2	AC107545	AC107545 Rattus no
7	709.6	56.7	2804	9	HSU73531	U73531 Human G pro
8	706.4	56.5	1923	9	BC033584	BC033584 Homo sapi
9	706.4	56.5	2238	6	AR049107	AR049107 Sequence
10	705.4	56.4	1953	6	AX549062	AX549062 Sequence
11	705.4	56.4	1953	6	HSY13248	Y13248 Homo sapien
12	704.8	56.3	1918	9	HSU73529	U73529 Human G pro
13	704.8	56.3	209772	9	AC099782	AC099782 Homo sapi
14	704.8	56.3	219553	2	HSX312887	AJ312887 Homo sapi
15	696	55.6	1032	9	AF124380	AF124380 Macaca mu
16	696	55.6	1032	9	AF237559	AF237559 Cercocebu
17	694.4	55.5	1072	9	AF291671	AF291671 Macaca fa
18	689.4	55.1	1029	6	AR158353	AR158353 Sequence
19	689.4	55.1	1029	6	AF007858	AF007858 Macaca ne
20	685	54.8	1037	6	AR158352	AR158352 Sequence
21	685	54.8	1037	9	AF007859	AF007859 Cercopith
22	679.2	54.3	1029	6	AR178364	AR178364 Sequence
23	679.2	54.3	1029	9	AF007545	AF007545 Homo sapi
24	679.2	54.3	1029	9	AF029759	AF029759 Homo sapi
25	679.2	54.3	1029	9	AF084229	AF084229 Pan trogl
26	679.2	54.3	1029	9	AY322543	AY322543 Homo sapi
27	679.2	54.3	1029	6	AX244235	AX244235 Sequence
28	676.8	54.1	1026	6	AR151483	AR151483 Sequence
29	676.8	54.1	1026	6	AX015624	AX015624 Sequence
30	648	51.8	1026	6	AR158351	AR158351 Sequence
31	304.8	24.4	449	9	AF354630	AF354630 Gorilla g
32	300.2	24.0	449	9	AF354631	AF354631 Pongo pyg
33	184.4	14.7	215176	2	BX546447	BX546447 Danio rer
34	183.4	14.7	1074	6	AR308572	AR308572 Sequence
35	183.4	14.7	1074	9	AY242127	AY242127 Homo sapi
36	183.4	14.7	1186	9	HSX13237	AJ132337 Homo sapi
37	183.4	14.7	2462	6	AX548955	AX548955 Sequence
38	183.4	14.7	2462	9	AF145440	AF145440 Homo sapi
39	183.4	14.7	2544	9	AF145439	AF145439 Homo sapi
40	183.4	14.7	2577	6	BD271525	BD271525 Anti-GPR-
41	183.4	14.7	2577	6	AR270532	AR270532 Sequence
42	183.4	14.7	2577	6	AR366325	AR366325 Sequence
43	183.4	14.7	2577	6	AX454125	AX454125 Sequence
44	183.4	14.7	2577	9	HSU45982	U45982 Human G pro
45	183.4	14.7	176968	9	AC005669	AC005669 Homo sapi

ALIGNMENTS

RESULT 1
AF301018
LOCUS AF301018 1251 bp mRNA linear ROD 25-NOV-2000
DEFINITION Mus musculus strain C57BL/6 chemokine receptor CXCR6 mRNA, complete cds.
ACCESSION AF301018
VERSION AF301018.1 GI:11342661
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1251)
AUTHORS Matloubian,M., David,A., Engel,S., Ryan,J.E. and Cyster,J.G.
TITLE A transmembrane CXC chemokine is a ligand for HIV-coreceptor Bonzo


```
QY 558 TGCTTGCTCATTTGGGTGGTCTCCCTGTTGGTTCCTTTGGCCACAGATCATCTATGGCCAT 617
Db 573 TGCTTGCTCATTTGGGTGGTCTCCCTGTTGGTTCCTTTGGCCACAGATCATCTATGGCCAT 514
QY 618 GTTCAAGATATTGACAGCTTATCTGTCAGTACCACAGTGGAGAGATATCCACTATGGTT 677
Db 513 GTTCAAGATATTGACAGCTTATCTGTCAGTACCACAGTGGAGAGATATCCACTATGGTT 454
QY 678 CTTGTTATACAGATGACTCTGGGGTCTTCTGCGCATGTCCTCATNTGATCTGTGCTAC 737
Db 453 CTTGTTATACAGATGACTCTGGGGTCTTCTGCGCATGTCCTCATNTGATCTGTGCTAC 394
QY 738 TCAGGCATTATCAAGACCTTGCTTCATGCTCGCAACTTCCAGAAGCAAACTCTCTAAAG 797
Db 393 TCAGGCATTATCAAGACCTTGCTTCATGCTCGCAACTTCCAGAAGCAAACTCTCTAAAG 334
QY 798 ATCATCTTCCTTTAGTGGCTGTGTCTGCTGTGACCCAGACACACCTTCAACCTTGGCATG 857
Db 333 ATCATCTTCCTTTAGTGGCTGTGTCTGCTGTGACCCAGACACACCTTCAACCTTGGCATG 274
QY 858 TTAATCCAAAGTACAAAGCTGGAGTACTATACCATACCAAGCTTTAAGTATGCCATGTA 917
Db 273 TTAATCCAAAGTACAAAGCTGGAGTACTATACCATACCAAGCTTTAAGTATGCCATGTA 214
QY 918 GTGACAGAGGCTATAGCATACTTTGGGGCTTGCCCTTAAACCTGTACTTTATGCGCTTTGTT 977
Db 213 GTGACAGAGGCTATAGCATACTTTGGGGCTTGCCCTTAAACCTGTACTTTATGCGCTTTGTT 154
QY 978 GCGTTAAAGTTCGGAAGAGCTGCTGGAACCTTATGAGGATATCGGCTGCTCTCCAC 1037
Db 153 GCGTTAAAGTTCGGAAGAGCTGCTGGAACCTTATGAGGATATCGGCTGCTCTCTCCAC 94
QY 1038 CTGGGAGTCTCAAGTCAATGGAAGCTTCTGAGGACAGTTCCTCAAGACTTGTTCGCTCC 1097
Db 93 CTGGGAGTCTCAAGTCAATGGAAGCTTCTGAGGACAGTTCCTCAAGACTTGTTCGCTCC 34
QY 1098 CACAATGTAGACACCAAGTATGTTCCAAATG 1130
Db 33 CACAATGTAGACACCAAGTATGTTCCAAATG 1

RESULT 5
AC126907 231959 bp DNA linear HTG 15-NOV-2002
LOCUS Rattus norvegicus clone CH230-150D22, WORKING DRAFT SEQUENCE, 6
DEFINITION unordered pieces.
ACCESSION AC126907
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 231959)
Muzny,D,Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Aman,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deranco,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Georgievski,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
```

```

Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Liu,J.,
Kowis,C., Kratt,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Loulseghe,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmood,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,S., Montanayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokelimeh,O., Okwuonu,G., Olarnungagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savary,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Snead,A., Sodergrat,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villabana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wlarczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,S., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 231959)
Worley,K.C.
Direct Submission
Submitted (10-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 231959)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Nov 15, 2002 this sequence version replaced gi:23195284.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GZSX
Center clone name: CH230-150D22
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 219672 bases at least Q40
Consensus quality: 211604 bases at least Q30
Consensus quality: 222280 bases at least Q20
Estimated insert size: 218775; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
table.
```


Chacko, J.J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregregios, E., Gear, K., Gill, R., Grady, M., Guerra, W., Guevara, M., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpachy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenzuewa, L., Loulsegged, H., Lozado, R. J., Lu, X., Ma, J., Maheshwar, M., Mahindaratne, M., Mahmood, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martineau, E., Manthey, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaekemelue, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Platter, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reich, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., S.J., Sanders, W., Savary, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steidle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, S., Thomas, D., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wlcek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

Direct Submission

Qy	434	AATGAACCTTACGTGTCATCTCACTCTCACTGTCATCACAGTGATGTTTCATTGT	493
Db	199725	AATGAACCTTACGTGTCATCTCACTCTCACTGTCATCACAGTGATGTTTCATTGT	199784
Qy	494	AGTGGTCCAGGCTACCAAGGCTTCAACCGCGCAGCTTAAGTGAAGATCTCGGGGCCAAAGT	553
Db	199785	AGTGGTCCAGGCTACCAAGGCTTCAACCGCGCAGCTTAAGTGAAGATCTCGGGGCCAAAGT	199844
Qy	554	CATTGCTTGTCTCATTTGGGTGCTCTCCCTGTGGTTCCTTTGGCCACAGATCATCTATGG	613
Db	199845	CATTGCTTGTCTCATCTGGGTGCTCTCCCTGTGGTTCCTTTGGCCACAGATCATCTACAG	199904
Qy	614	CCATGTTCAAGATATTCACAAAGCTTATCTGTGCTACACAGTACAGGAGATATCCACTAT	673
Db	199905	CAGATTTGGACATATTCACAGGCTTATCTGTGCTACACAGGAGATATCGACTGT	199964
Qy	674	GGTTCTTGTATACAGATGACTCTGGGGTCTTCTCTGCCATTTGCTCACTATGATCTGTG	733
Db	199965	GGTTCTTGTCTACAGATGACTCTGGGGTCTTCTCTGCCATTTGCTCACTATGATCTGTG	200024
Qy	734	CTACTCAGGCATATTCAGACCTTCTTCTCATGCTCGAACTTCCAGAGCACAAATCTCT	793
Db	200025	CTACTCAGGCATATTCAGACCTTCTTCTCATGCTCGAACTTCCAGAGCACAAATCTCT	200084
Qy	794	AAAGATCATCTTCTTGTAGTGGCTGTCTTCTGCTGACCCAGACACCCCTTCAACCTTGC	853
Db	200085	AAAGATCATCTTCTTGTAGTGGCTGTCTTCTGCTGACCCAGACACCCCTTCAACCTTGC	200144
Qy	854	CATGTTAATCCAAAGTACAGCTGGGAGTACTATACCAATACCAAGCTTTAAGTATGCCAT	913
Db	200145	CATGTTAATCCAGTGAATAGTTGGGAATACAGTACCAATACCAAGCTTTAAGTATGCCAT	200204
Qy	914	CGTAGTGACAGAGGCTATAGCATACTTTCGGGCTTGCCTTAAACCTGTACTTTATGCCCTT	973
Db	200205	CTTAGTGACAGAGGCTATAGCATACTTTCGGGCTTGCCTTAAACCTGTACTTTATGCCCTT	200264
Qy	974	TGTTGGCTTAAAGTTCGGGAAGACGCTCTGGAACCTTATGAAGATATCGGCTGCTCTC	1033
Db	200265	TGTTGGCTTAAAGTTCGGGAAGACCTCTGGAACCTTATGAAGATATCGGCTGCTCTC	200324
Qy	1034	TCACCTGGGAGTCTCAAGTCAATGGAAGTCTCTGAGGACAGTTCACAGACTTGTCTCG	1093
Db	200325	TCGCTGGGAGTCTCAAGTCAATGGAAGTCTCTGAGGACAGTTCACAGACTTGTCTCG	200384
Qy	1094	CTCCCAATGTAGAGACCAACAGTATGTTCCAAATTTAGTAGGCTTCCACACTTAGA	1153
Db	200385	CTCCCAATGTAGAGACCAACAGTATGTTCCAAATTTAGTAGGCTTCCACACTTAGA	200444
Qy	1154	GAAGTTAATACAGAAATTC-TAGGACATGGCTGTATCATTTGGATGCAACAGAAAGC	1212
Db	200445	GAAGTTAATACAGAAATTC-TAGGACATGGCTGTATCATTTGGATGCAACAGAAAGC	200504
Qy	1213	TTTGCTTATAGCATGTGGAGTATCATGGAGAGCC	1247
Db	200505	TTTGCTTATAGCATGTGGAGTATCATGGAGAGTC	200539
RESULT 7			
HSU73531			
LOCUS			
DEFINITION			
Human G protein-coupled receptor STRL33.3 (STRL33) mRNA, complete cds.			
ACCESSION			
U73531			
VERSION			
U73531.1			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1 (bases 1 to 2804)			
Liao,F., Alkhatib,G., Peden,K.W., Sharma,G., Berger,E.A. and Farber,J.M.			
TITLE			
STRL33, A novel chemokine receptor-like protein, functions as a fusion cofactor for both macrophage-tropic and T cell line-tropic HIV-1			
J. Exp. Med. 185 (11), 2015-2023 (1997)			
MEDLINE			
97311099			
PUBMED			
9166430			
REFERENCE			
2 (bases 1 to 2804)			
Liao,F. and Farber,J.M.			
Direct Submission			
Submitted (07-OCT-1996) Laboratory of Clinical Investigation,			
National Institute of Allergy and Infectious Diseases, NIH, 9000			
Rockville Pike, Bldg 10, Rm 11N-228, Bethesda, MD 20892, USA			
JOURNAL			
Location/Qualifiers			
FEATURES			
source			
1..2804			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosome="3"			
/note="expressed in T cells"			
1..2804			
/gene="STRL33"			
1463..2491			
/gene="STRL33"			
/note="contains allelic differences compared to STRL33.1 encoded by GenBank Accession Number U73529; contains seven transmembrane domains"			
/codon_start=1			
/product="G protein-coupled receptor STRL33.3"			
/protein_id="AAB61457.1"			
/db_xref="GI:2209288"			
/translation="MAEHDYEDYFSSFNDSOEHQAFLOFSKVPCLPMYLVFVFC			
GLGNSLVLVISLTFYHKLQSLTDVFLNVLADLVFVCTLPFWAYAGIHEWFGQVMC			
KSLGIYTFINFTSMILTCITDFTFVVKATKAYNQAKRTWQKVTSLLIWVLSL			
LVSPQIYGNVFNLDKLCIGYDEALSTVLTATMTLGFLLPLTITVVCYVIKTL			
LHAGGPKHRSKATILFLMAVFLITOMPENLMKFLISTHWEYVAMSFHYTIWTERAI			
AYLRACINPVLVAFVSLKPRNFVKLVKDGCLPILGVSHQWKSSEDNSKTFSSAHNV			
EATSNFQL"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches			
Score 709.6; DB 9; Length 2804;			
78.8%; Pred. No. 1.2e-209;			
Conservative			
0; Mismatches 229; Indels 6; Gaps 2;			
Qy	141	TTCAACAAATTCACGTATACAGCC---AGAGAACAAACGCTTCCTAAAGTTCAGAGGAG	197
Db	1496	TTCAACAAATTCACGTATACAGCC---AGAGAACAAACGCTTCCTAAAGTTCAGAGGAG	1555
Qy	198	GTCTTTTGGCCCTGTGTACCTGCTAGTGTCTTCTTCTGCTCTGCTAGGAACTCCCTG	257
Db	1556	GTCTTTTGGCCCTGTGTACCTGCTAGTGTCTTCTTCTGCTCTGCTAGGAACTCCCTG	1615
Qy	258	GTTCTGATATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTGTTTCTGCTG	317
Db	1616	GTGCTGCTCATATCCATCTCTACCAATAGTTCAGAGCCCTGACGATGTGTTCTGCTG	1675
Qy	318	AACCTGCCCCCTGCTGACCTGCTGTGTTGCTGTACTCTGCCCTTTGGGCTATGAGGC	377
Db	1676	AACCTGCCCCCTGCTGACCTGCTGTGTTGCTGTACTCTGCCCTTTGGGCTATGAGGC	1735
Qy	378	ACCTATGAGTGGGTCTTTGGCAGATCATGTGCAAAACTCTTCGAGGCATGTATACAATG	437
Db	1736	ATCCATGAATGGGTGTTTGGCCAGGTCTATGTGCAAAAGCTTACTGGGCATCTACATATT	1795
Qy	438	AACCTTCTACGTGTCCTGCTCACTCTCACTGATCATGATGATGATGATGATGATGATG	497
Db	1796	AACCTTCTACGTGTCCTGCTCACTCTCACTGATCATGATGATGATGATGATGATGATG	1855
Qy	498	GTCAGAGCTACCAAGGCTTCAACCGCAGGCTTAAGTGAAGATCTGGGGCCCAAGTCATT	557
Db	1856	GTTAAGGCCCAAGGCTTCAACCGCAGGCTTAAGTGAAGATGATGATGATGATGATGATG	1915
Qy	558	TGCTTGTCTCATTTGGGTGGTCTCCCTGTTGGTTCCTTGGCCACAGATCATCTATGCCAT	617

Db 1916 AGCTTGCTCATCTGGGTGATATCCCTGCTGGTTCCTTGCCCAAAATTAATATGGAAT 1975
 QY 618 GTTCAAGATATGCAAGCTTATCTGTAGTACCACTGAGGAGATATCCACTATGTT 677
 Db 1976 GCTTTAATCTCGAAGCTCATATGTGGTACCACTGAGGCAATTTCCACTGTGTT 2035
 QY 678 CTTGTATACAGATCACTCTGGGGTCTTCTCTGCAATGCTCTCACTATGATCTGTGCTAC 737
 Db 2036 CTTGCCACCAAGATGACACTGGGGTCTTCTTGCCCACTGCTCACTATGATCTGTGCTAT 2095
 QY 738 TCAGCAATATCAAGACCTTGTCTCATGCTGCAAACTTCCAGAGCACAATCTCTAAG 797
 Db 2096 TCAGTCATAATCAAAACACTGCTTATGCTGAGGCTTCCAGAGCAGATCTCTAAG 2155
 QY 798 ATCATCTTCTTGTAGTGGCTGTGTTCTGCTGACCCAGACACCCCTTCAACCTTGCCATG 857
 Db 2156 ATCATCTTCTGCTGATGCTGTGTTCTGCTGACCCAGATGCTTCAACCTCATGAAG 2215
 QY 858 TTAATCCAAAGTACAACTGGGAGTACTATACCAATCAAGCTTTAAGTATGCCATGTA 917
 Db 2216 TTCAATCCGAGCACACTGGGAATACTATGCTCATGACCAAGCTTTCACTACCACTCATG 2275
 QY 918 GTGACAGAGGCTATAGCATCTTTCGGGCTTGCCTTAACCTGACTTATGCTTCTGTTT 977
 Db 2276 GTGACAGAGGCTATGCACTACCTGAGGCTGCTTAACTGCTCTATGCTTCTGCT 2335
 QY 978 GGCCTTAAAGTCCGAAAGAACTCTGGAACCTTATGAAGGATATCGGCTGCTCTCTCAC 1037
 Db 2336 AGCCTGAAGTTTCGAAAGAACTCTGGAACCTTGTGAAGGACATTTGCTGCTCTCTCAC 2395
 QY 1038 CTGGGAGTCTCAAGTCAATGGAAGTCTCTGAGGACAGTCCAGACCTTGTCTGCTCTCC 1097
 Db 2396 CTTGGGCTCTCACTCAATGGAATCTCTGAGGACAAATTCAGACCTTTTCTGCTCTCC 2455
 QY 1098 CACAATGTAGACACCACTAGTATGTTTCAATTTAGTAGGCTTCCACCTTCCACCTAGAGAG 1157
 Db 2456 CACAATGTGGAGCCACCACTAGTTTCA--GTTATAGGCTTCCAGGGTTTCGAAAG 2512
 QY 1158 TTAATAACAGATTTAGAGAGATGCTGTATCATTTGATGTCACACAGAAAGCTTTCG 1217
 Db 2513 ACTGCTCTGGAATTTGCAAGGATGCTGTGCTCTTGTATGTTGATGAGGAGGCTTGT 2572
 QY 1218 TTATAGCATGTGGAGTATCATGAGAG 1245
 Db 2573 TTATAGCTTGCATCTCATGAGAG 2600
 RESULT 8
 BC033584
 LOCUS
 DEFINITION Homo sapiens chemokine (C-X-C motif) receptor 5, mRNA (cdna clone MGC:34420 IMAGE:5225727), complete cds.
 ACCESSION BC033584
 VERSION BC033584.1 GI:21707372
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1923)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diachenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Donald,M.F., Casavant,T.L., Schetz,T.B., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,

Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Buffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Bouffard,M.C., Rodriguez,A.C., Grimoc,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E., Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A. full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22389257
 MEDLINE
 PUBMED 12477932
 2 (bases 1 to 1923)
 Strausberg,R.
 Direct Submission
 TITLE
 JOURNAL
 REMARK
 COMMENT
 Contact: MGC help desk
 Email: gcaps-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc_mgc@hri.nih.gov
 Akhtar,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masello,C., Maekari,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 49 Row: n column: 21
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 5730105.
 FEATURES
 source
 1..1923
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:34420 IMAGE:5225727"
 /tissue_type="Pancreas, Spleen, adult pooled"
 /clone_lib="NIH MGC 120"
 /lab_host="DH10E"
 /notes="Vector: pCMV-SPORT6"
 gene
 1..1923
 /genes="CXCR6"
 /notes="synonyms: STRL33, TYMSTR, BONZO"
 /db_xref="LocusID:10663"
 /db_xref="MIM:605163"
 33..1061
 /codon_start=1
 /product="G protein-coupled receptor TYMSTR"
 /protein_id="AAH33584.1"
 /db_xref="GI:21707373"
 /db_xref="LocusID:10663"
 /translations="MAEDHYEDYGFSSFNDSQEEHQDFLOPSKVFPCMLVWVFC GLVNSIVLVSIFVHKLOSLTDVFLVNLPLADLVVCTLPFWAYAGIHEWVFGVMC KSLIGYITNFYFMSLILICITVDRIYVVKATKAYNQAKQKNTGKVTSLLLVLSL LVSLPQIYGNVFNLDKLCIGYDEAISTVLAQTQTLGFLLEITWICYSVIITKL THGAGQKRSKLIIFLVAVFLTLQTPNLMKFIKSTHWEYIYATMSFHTIMVTAI AYLRACINPVLYAFVSLKFRKFWKLVDGICLPYLGVSHQWKSSEDSNKTFSASHV EATSMFOL"
 174..896
 /notes="7tm_1; Region: 7 transmembrane receptor (rhodopsin family)"
 misc_feature

[illegible]

Db 905 CTTGCCACCCAGATGACACTGGGGTCTTCTTCCACTGCTCACCATGATTGCTGCTAT 964
Qy 738 TCAGGCATTATCAAGACCTTCTTCATGCTCGAACTTCCAGAGCACAATCTCTAAAG 797
Db 965 TCAGTCATATCAAAACACTGCTTCAAGCTGGAGGCTTCCAGAGCACAAGATCTCTAAAG 1024
Qy 798 ATCATCTTCTGCTAGTGGTCTGCTGCTGACCCAGACACACCTTCAACCTTGGCATG 857
Db 1025 ATCATCTTCTGCTAGTGGTCTGCTGCTGACCCAGATGCCCTTCAACCTCATGAAG 1084
Qy 858 TTAATCCAAAGTACAGCTGGGAGTACTATACCATTAACCAAGCTTTAAGTATGCCATCGTA 917
Db 1085 TTTCATCCGAGACACACCTGGGAATATATGCGCATGACCAAGCTTCACTACACCATCATG 1144
Qy 918 GTGACAGAGCTATAGCATACTTTCGGGCTTGGCTTAAACCTGTACTTTATGCCCTTGT 977
Db 1145 GTGACAGAGCCATGTCATACCTGAGGGCTTGCCTTAAACCTGTGCTCTATGCCCTTGT 1204
Qy 978 GGCCTTAAAGTTCGGGAAGACGCTCTGGAACTTATGAAGATATGCGCTGCTCTCTCAC 1037
Db 1205 AGCCTGAAGTTTCGAAAGAACTTCTGAAACTTGTGAAGGACATTTGCTGCTCCCTTAC 1264
Qy 1038 CTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCAGACTTGTCTGCTCC 1097
Db 1265 CTGGGCTCTCATCATCAATGGAATCTTCTGAGGACAAATCCAGACTTTTCTGCTCC 1324
Qy 1098 CACATGTAGAGACACCAAGTATGTTCCAAATGTATAGGCTTCCACACTTACAGAG 1157
Db 1325 CACAATGTGGAGGCCACCAAGCATGTTCCTCA--GTTATAGGCTTCCAGGGTTTCGAGAA 1381
Qy 1158 TTAATAACAGAACTTCTAGGAGCATGCTGATCAATTTGGATGACACAGAAAGCTTTGC 1217
Db 1382 GCTGCTCGAAATTTGCAAGGATGCTGCTGCTGCTTGTATGTGTGAGGAGGCTTTGT 1441
Qy 1218 TTATAGCATGTGGAGTATCATGAGAAG 1245
Db 1442 TTATAGCTTGGCAATCTCATGAGAAG 1469

RESULT 10
AX549062
LOCUS AX549062 1953 bp DNA linear PAT 26-NOV-2002
DEFINITION Sequence 347 from Patent WO02061087.
ACCESSION AX549062
VERSION AX549062.1 GI:25813841
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Burmer,G.C., Roush,C.L. and Brown,J.P.
TITLE Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides
JOURNAL Patent: WO 02061087-A 347 08-AUG-2002;
Lifespan Biosciences, Inc. (US)
FEATURES
source 1..1953
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 56.4%; Score 705.4; DB 6; Length 1953;
Best Local Similarity 78.5%; Pred. No. 2.4e-208;
Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;
Qy 141 TTCACAAATTCAGTGTATACAGCC--AGGAGAACAAACCTTCCCTAAAGTTCAAGGAG 197
Db 115 TTCAGCAGTTTCAATGACAGAGCCAGGAGGAGCATCAGACTTCTGCGAGTTCCAGCAAG 174
Qy 198 GTCCTTTTGGCCCTGTGTACTGTGTGTTGCTTTGGACTGTAGGAAACTCCCTG 257

RESULT 11

Db 175 GTCTTTCTGCCCTGCATGTACTGTTGTTGTTCTGTGTTCTGTTGGGAATCTCTGT 234
Qy 258 GTTCTGATTATATACATTTCTACCAGAAGCTGAGGACTCTGCAGAGATGTGTTTCTGTG 317
Db 235 GTGCTGTGCATATCCATCTTCTACCATAAAGTTGCAGAGCCTGACGGAATGTGTTCTG 294
Qy 318 AACTTGGCCCTGAGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 377
Db 295 AACCTACCCCTGAGCTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 354
Qy 378 ACCTATGAGTGGGCTTCTTGGCACAGTCACTGTGCAAACTCTTTCGAGGCATGTATACA 437
Db 355 ATCCATGAATGGGTTTGGCCAGTCACTGTGCAAGAGCTACTTGGGCATCTACACTATT 414
Qy 438 AACTTCTACGTGTCATGCTCCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
Db 415 AACTTCTACAGTCCATGCTCATCTCCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
Qy 498 GTCCAGGCTACCAAGGCTTCAACCGGCGAGCTCAAGTGGAGAGATCTCGGGCCCAAGTCA 557
Db 475 GTTAAGGCCCAAGGCTCAACACGACGACGACGACGACGACGACGACGACGACGACGAC 534
Qy 558 TGCTTGTCTATTTGGTGTGCTCCTCTGTTGTTCTTGTGCTGCTGCTGCTGCTGCTGCTG 617
Db 535 AGCTTGTCTCATCTGGGATATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 594
Qy 618 GTTCAAGATATGACAGCTTATCTGTCAGTACACAGTGGAGAGATATCCACTATGCTG 677
Db 595 GTCTTTAATCTGCAAGCTCATATGTTGTTTCCATGACGAGGCAATTTCCACTGTGTT 654
Qy 678 CTGTTTATACAGATGACTCTGGGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
Db 655 CTGTCACCCAGATGACACTGGGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
Qy 738 TCAGGCATATCAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 797
Db 715 TCAGTCATAATCAAAACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774
Qy 798 ATCATCTTCTCTGTAGTGGCTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857
Db 775 ATCATCTTCTGCTGATGGCTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
Qy 858 TTAATCCAAAGTCAAGCTGGAGTACTATACCAATACCAAGCTTTAAGTATGCCATGTA 917
Db 835 TTCATCCGAGCACACACTGGGAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
Qy 918 GTGACAGAGCTATAGCATCTTTCGGGCTTGGCTTAAACCTGCTGCTGCTGCTGCTGCTG 977
Db 895 GTGACAGAGCCATGCGATACCTGAGGCTGCTGCTTAAACCTGCTGCTGCTGCTGCTG 954
Qy 978 GGCTTAAAGTTCGGAAGAACGTTCTGGAACCTTATGAAGGATATCGGCTGCTGCTGCTG 1037
Db 955 AGCCTGAAGTTTCGAAAGAACTTCTGGAACCTTGTGAAGGACATTTGTTGCTGCTGCT 1014
Qy 1038 CTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCAGACTTGTGTTGCTGCT 1097
Db 1015 CTGGGAGTCTCAATCAATGGAATCTTCTGAGGACAAATTCGAGGCAATTTCTGCTGCT 1074
Qy 1098 CACAATGTAGAGACCAACCAAGTATGTTTCCAAATGTAGTAGGCTTGGCCACTTTCAGAG 1157
Db 1075 CACAATGTGAGGCCACCAAGCATGTTTCCA--GTTATAGGCTTGGCCAGGTTTCAGAA 1131
Qy 1158 TTAATACAGAAATCTAGGAGCATGCTGATCATTTGGATGCAACAGAAAGCTTTGCTG 1217
Db 1132 GCTGCTCTGGAATTTGCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1191
Qy 1218 TTATAGCATGTGGAGTATCATGAGAAG 1245
Db 1192 TTATAGCTTGGCAATCTCATGAGAAG 1219

REFERENCE 1 (bases 1 to 1918)
 AUTHORS Liao,F., Alkhatib,G., Peden,K.W., Sharma,G., Berger,E.A. and Farber,J.M.
 TITLE STRL33, A novel chemokine receptor-like protein, functions as a fusion cofactor for both macrophage-tropic and T cell line-tropic HIV-1
 J. Exp. Med. 185 (11), 2015-2023 (1997)
 J97311099
 MEDLINE 9166430
 PUBMED
 REFERENCE 2 (bases 1 to 1918)
 AUTHORS Liao,F. and Farber,J.M.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1996) Laboratory of Clinical Investigation, National Institute of Allergy and Infectious Diseases, NIH, 9000 Rockville Pike, Bldg 10, Rm 11N-228, Bethesda, MD 20892, USA
 FEATURES
 Location/Qualifiers
 1..1918
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosomes="3"
 /cell_type="lymphocytes"
 /note="expressed in tumor infiltrating lymphocytes"
 1..1918
 /gene="STRL33"
 31..1059
 /gene="STRL33"
 /note="contains seven transmembrane domains"
 /codon_start=1
 /product="G protein-coupled receptor STRL33.1"
 /protein_id="AB61456.1"
 /db_xref="GI:2209285"
 /translation="MAEDHYDEYGFSSNDSEHQDFLQSKVFLPCMVLVVFVCGLVGNLSLVIVSIFVYHKLQSLTDVFLVNLPLADLVFVCTLPFWAYAGHEWVFGVMSKSLGIYINFTISMLILCTVDVFIIVVWVATKAYNOOAKSMWTKVTSLLIWLISLVLPQIYGVNFDLKLIGVHDEAIPTVLATQMTGFLPLLLTIVCYSLIKTL LHAGSFQKHSRLKILFLYMAVELLQMPFLMKFIRSHWYAYTSEHYIMTEALAYLRACLPLVLYAFSLFRKNFWLVDIGLPLVGVSHQWKSSEDSNKTFSASHNV EATSMFQL"
 ORIGIN
 Query Match 56.3%; Score 704.8; DB 9; Length 1918;
 Best Local Similarity 78.5%; Pred. No. 3.6e-208;
 Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;
 141 TTCACCAATTCAGTGATAACAGCC---AGGAGAACAAACGCTTCCTAAAGTTCAAGGAG 197
 DB |||||
 64 TTCAGCAGTTTCATGACAGCAGCCAGGAGGAGCATCAAGACTTCCTGCAGTTCAGCAAG 123
 QY |||||
 198 GTCTTTTGGCCCTGTGTACTGT 257
 DB |||||
 124 GTCTTTTGGCCCTGTGTACTGT 183
 QY |||||
 258 GTTCTGATTATATACATTTTCTACCAAGCTGTAGGACTCTGACAGATGTGTGTGTGTGTGTGT 317
 DB |||||
 184 GTGCTGGTTCATATCCATCTTACCATTAAGTTGCAGAGCCTGACGGAATGTCTCTGGTG 243
 QY |||||
 318 AACTTGGCCCTGGCTGACTGT 377
 DB |||||
 244 AACCTACCCCTGGCTGACTGT 303
 QY |||||
 378 ACCTATGAGTGGTCTTTGGCAGCAGTCTATGTGTGCAAAACTCTTCGAGGCACTATACAAATG 437
 DB |||||
 304 ATCCATGAATGGGTGTGTGGCCAGGTCTATGTGCAGAGCCTACTGGGCACTACACTATT 363
 QY |||||
 438 AACTTCTAGTGTCCAGCTGCTCACTCTCACTGCATCAGTGGATCGTTTCATTGTAGTG 497
 DB |||||
 364 AACTTCTACAGTCCATCTCTCATCTCACTGCATCAGTGGATCGTTTCATTGTAGTG 423
 QY |||||
 498 GTCCAGGCTACCAAGGCTTCAACCGCAGGCTAAGTGGAGACTCGGCCCAAGTCAATT 557
 DB |||||
 424 GTTAAGGCCACCAAGGCTTACACGACCAAGGAGGATGCTCGGGCAAGGTCAAC 483

QY 558 TGCTTGCTCATTTGGTGGTCTCCTGT 617
 DB |||||
 484 AGCTTGCTCATCTGGTGTGATATCCTGT 543
 QY |||||
 618 GTTCAAGATATTGCAAGCTTATCTGTCAGTACCACAGTACAGGAGATATCCACTATGTTT 677
 DB |||||
 544 GTCTTTAATCTCGACAGCTCATATGT 603
 QY |||||
 678 CTTGTTATACAGATGACTCTGGGTTCTTCTGTCCATTGTCTCACTATGATTCTGTGTCTAC 737
 DB |||||
 604 CTTGCCACCCAGATGACACTGGGGTCTCTTCTGTGCACTGTCTCACCATTGTCTGTCTAT 663
 QY |||||
 738 TCAGGCAATTAACAAGCTTCTGCTCATGCTCGAACTTCCAGAGGACAAATCTCTTAAG 797
 DB |||||
 664 TCAGTCATATCAAAACACTGCTTCATGCTGGAGGCTTCCAGAGGACAGATCTCTTAAG 723
 QY |||||
 798 ATCATCTTCTCTGTAGTGGCTGTCTTCTGTGACCCAGACACCCCTTCAACCTTGCCTATG 857
 DB |||||
 724 ATCATCTTCTCTGT 783
 QY |||||
 858 TTAATCCAAAGTACAGCTGGAGTACTATACCAATACCAAGCTTTAAAGTATGCCATCGTA 917
 DB |||||
 784 TTATCCGCGACACACACTGGGAATACTATGCCATGACCAAGCTTTCACTACACCATCATG 843
 QY |||||
 918 GTGACAGAGGCTATAGCATACTTTCGGGCTTGCTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 977
 DB |||||
 844 GTGACAGAGGCTATCGCATACCTGAGGGCTGCTTAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 903
 QY |||||
 978 GCGTTAAAGTTCGGAAGAACGCTGTGAAACTTATGAAGGATATCGGCTGCCCTCTCTCTAC 1037
 DB |||||
 904 AGCCTGAAGTTTCGAAAGAACTTCTGAAAGCTTGTGAAGGACATTTGTTGCCCTCCCTTAC 963
 QY |||||
 1038 CTGGGAGTCTCAAGTCAATGAAAGTCTTCTGAGGACAGTTCCAGAGCTTGTCTGTGCTCC 1097
 DB |||||
 964 CTTGGGCTCTCATCATCAATGAAATCTTCTGAGGACAAATTCAGAGACTTTTCTGCTCC 1023
 QY |||||
 1098 CACAATGTAGAGACCAAGCATGTGTTCAAATTTAGTAGGCTTGGCACACTTTAGAGAAG 1157
 DB |||||
 1024 CACAATGTGAGGCCACCAAGCATGTGTCCA---GTTATAGGCTTGCAGGGTTTCGAGAA 1080
 QY |||||
 1158 TTAATACAAATCTAGGAGCATGCTGTATCATTTGGATGCAACAGAAAGCTTTGCT 1217
 DB |||||
 1081 GCTGCTCTGAAATTTGCAAGTCAAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1140
 QY |||||
 1218 TTATAGCATGTGAGTATCATGGAAG 1245
 DB |||||
 1141 TTATAGCTTGGCATCTCTCATGGAAG 1168

RESULT 13
 AC099782 . 209772 bp DNA linear PRI 23-APR-2002
 LOCUS Homo sapiens chromosome 3 clone RP11-852E15, complete sequence.
 DEFINITION AC099782 AC024150
 ACCESSION AC099782.2 GI:20270170
 VERSION HTG.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 209772)
 AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 209772)
 AUTHORS Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2001) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA
 REFERENCE 3 (bases 1 to 209772)
 AUTHORS Kaul,R.K., Olson,M.V., Zhou,Y., James,R.A., Rouse,G., Wu,Z., Saenphimmachak,C., Phelps,K.A., Raymond,C. and Haugen,E.D.

TITLE Direct Submission
 JOURNAL Submitted (23-APR-2002) Genome Center, University of Washington,
 Box 352145, Seattle, WA 98195, USA
 COMMENT On Apr 23, 2002 this sequence version replaced gi:17027297.

----- Genome Center
 Center: University of Washington Genome Center
 Center Code: UWGC
 Web site: <http://www.genome.washington.edu>
 Contact: uwgchgs@u.washington.edu
 Drafting Center: BCM
 ----- Project Information
 Center project name: chr-3
 Center clone name: RP11-852E15 (bc0572)
 ----- Summary Statistics
 Sequencing vector: M13; L08821; 56% of reads
 Chemistry: Dye-terminator; 44% of reads
 Chemistry: Dye-terminator; 37% of reads
 Chemistry: Dye-terminator; 8% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 209522 bases at least Q40
 Consensus quality: 209746 bases at least Q30
 Consensus quality: 209771 bases at least Q20
 Insert size: 209772; sum-of-contigs
 Quality coverage: 8.2x in Q20 bases; sum-of-contigs

 Overlapping Sequences:
 5': RP11-165I16 (UWGC:bc0278) AC005669
 3': RP11-91B8 (UWGC:bc0216) AC026349

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

ECORI		BglII		HindIII	
SeqDerMap	EngPrnt	SeqDerMap	EngPrnt	SeqDerMap	EngPrnt
8696	8835	15519	16344	2966	2960
6	<800	2067	2076	6382	6486
14153	14070	9563	9576	512	<800
3038	3060	1858	1823	449	<800

6848	6876	1282	1264	1805	1734
440	<800	462	<800	3678	3815
7429	7388	4203	4247	437	<800
5315	5487	199	<800	1567	1549
3186	3234	4445	4430	346	<800
6959	6876	16180	16344	1438	1415
3159	3149	8783	8725	3415	3413
2762	2727	3192	3400	9698	9638
5886	5841	6244	6165	1543	1549
1745	1725	604	<800	3378	3413
1551	1615	303	<800	2332	2349
1871	1901	7626	7540	520	<800
1186	1181	2931	2969	1753	1734
5076	5089	7121	7139	4580	4573
10486	10324	919	827	3707	3815
646	<800	1495	1472	2852	2843
8906	8835	3686	3694	5743	5750
199	<800	2174	2250	2874	2862
237	<800	2842	2969	2360	2196
7120	7162	1800	1823	2376	2543
7093	7162	791	<800	651	<800
11716	11593	2269	2250	2826	2862
1643	1615	2567	2620	1219	1209
2201	2206	560	<800	2061	2043
7707	7749	1048	1051	1601	1549
3648	3594	5744	5750	6291	6322
7397	7388	9748	9576	12244	12340
16603	16829	4114	4052	6104	6091
1959	1901	3370	3400	3840	3815
414	<800	2921	2969	31	<800
1680	1615	196	<800	4833	4759
754	766	2080	2076	2973	2960
1083	1084	1600	1596	9069	9129
5961	5841	3638	3694	2030	2043
10484	10324	374	<800	2041	2043
5470	5487	4256	4247	13	<800
5665	5841	11051	10673	2010	2043

	-----	5185	5315	3021	3222	578	6091
	-----	130	<800	1804	1823	1108	1110
	-----	2174	2206	16809	16344	13385	13422
	-----	613	<800	4011	4052	1572	1549
	-----	279	<800	658	<800	2128	2196
	-----	4723	4676	1505	1472	265	<800
	-----	45	<800	10543	10224	1011	1014
	-----	6943	6876	7521	7540	415	<800
	-----			59	<800	4801	4759
	-----			4550	4430	5104	4953
	-----			110	<800	829	862
	-----			6055	5997	3842	3815
	-----					3117	3089
	-----					2125	2196
	-----					8046	7986
	-----					4498	4461

Query Match 56.3%; Score 704.8; DB 9; Length 209772;
 Best Local Similarity 78.5%; Pred. No. 6.6e-208;
 Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;

Qy	141	TTCAACRAATTCAGTGATAACGAGCC--AGGAGAACAAACGGTCTCCTAAAGTTCGAAGG	197
Dd	132440	TTCAGCATTTCAATGACAGCGCCAGGAGGAGCATCAAGACTTCTCGAGTTCAGCAAG	132499
Qy	198	GTCCTTTTTGCCCTGTGTACTGGTAGTGTGTTGTTCTTGGA	257
Dd	132500	GTCATTCTGCCCTGCATGTACTGGTGTGTTGTTCTGTGGTCTGGTGGGAACTCTCTG	132559
Qy	258	GTTCTGATTATATACATTTTCTACAGAGCTGAGGACTCTGACAGATGTGTTCTGCTG	317
Dd	132560	GTGCTGGTCATATCCATCTTCTACCATAAGTTGCAGAGCCTCACGGATGTGTTCTGGTG	132619
Qy	318	AAC TTGCCCTGGCTGACCTGGTGTGTTGTTCTGTACTCTGCCCTTTTGGSCCTATGCAGC	377
Dd	132620	AACCTACCCCTGGCTGACCTGGTGTGTTGTTCTGCACTCTGCCCTTCTGGSCCTATGCAGC	132679
Qy	378	ACCTATGAGTGGTCTTTGGCACAGTCATGTGCCAAA	437
Dd	132680	ATCCATGAATGGGTGTTTGGCCAGGTCATGTGCCAAGCCTACTGGGCATCTACACTATT	132739
Qy	438	AAC TTCTACGTGTCATGCTCACTCTCACTCGATCAGTGGATCGTTTTCATTGTAGTG	497
Dd	132740	AAC TTCTACAGTCCATGCTCATCTCCTCACTGATCAGTGGATCGTTTTCATTGTAGTG	132799
Qy	498	GTCAGGCTACAAGGCCCTTCAACGGCAGGCTTAAGTGAAGATCTGGGGCCAAGTCATT	557
Dd	132800	GTTAAGGGCCACAAGGCCCTACACCAGCAAGGATGACCTGGGGCAAGTCCACC	132859
Qy	558	TGCTTGCTCATTTGGGTGGTCTCCCTGTTTGGTTTCTTTSCCAGAGATCATCTATGSCCAT	617
Dd	132860	AGCTTGCTCATCTGGGTGATATCCCTGTGTTTCTTGCCCCAAATATCATATGGCAAT	132919
Qy	618	GTTCAAGATATGCAAGCTTANCTGTGAGTACCAAGTGGAGGATATCCACTATGGTT	677
Dd	132920	GTC TTTTAACTCGAAGAAGCTCATATGTGTTTACCATGACGAGGCAATTTCCACTGTGGTT	132979

represented
 by 100 Ns. Contig 1: 1-4409 bp Contig 2: 4510-5012 bp Contig 3:
 5113-9787 bp Contig 4: 9888-11842 bp Contig 5: 11943-14635 bp
 Contig 6:
 14736-25152 bp Contig 7: 25253-34186 bp Contig 8: 34287-37648 bp
 Contig
 9: 37749-39128 bp Contig 10: 39229-56598 bp Contig 11:
 56699-62191 bp
 Contig 12: 62292-71083 bp Contig 13: 71184-81023 bp Contig 14:
 81124-110809 bp Contig 15: 110910-113803 bp Contig 16:
 113904-131104 bp
 Contig 17: 131205-134372 bp Contig 18: 134473-136524 bp Contig
 19:
 136625-140189 bp Contig 20: 140290-148706 bp Contig 21:
 148807-149537 bp
 Contig 22: 149638-169311 bp Contig 23: 169412-219553 bp.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 23 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * been provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 4409: contig of 4409 bp in length
 * 4410 4509: gap of 100 bp
 * 4510 5012: contig of 503 bp in length
 * 5013 5112: gap of 100 bp
 * 5113 9787: contig of 4675 bp in length
 * 9788 9887: gap of 100 bp
 * 9888 11842: contig of 1955 bp in length
 * 11843 11943: gap of 100 bp
 * 11943 14635: contig of 2693 bp in length
 * 14636 14736: gap of 100 bp
 * 14736 25152: contig of 10417 bp in length
 * 25153 25253: gap of 100 bp
 * 25253 34186: contig of 8934 bp in length
 * 34187 34287: gap of 100 bp
 * 34287 37649: contig of 3362 bp in length
 * 37649 37749: gap of 100 bp
 * 37749 39128: contig of 1380 bp in length
 * 39129 39229: gap of 100 bp
 * 39229 56599: contig of 17370 bp in length
 * 56599 56699: gap of 100 bp
 * 56699 62191: contig of 5493 bp in length
 * 62192 62292: gap of 100 bp
 * 62292 71083: contig of 8792 bp in length
 * 71084 81023: contig of 9840 bp in length
 * 81024 81123: gap of 100 bp
 * 81124 110809: contig of 29686 bp in length
 * 110810 113803: contig of 2894 bp in length
 * 113804 113904: gap of 100 bp
 * 113904 131104: contig of 17201 bp in length
 * 131105 131204: gap of 100 bp
 * 131205 134372: contig of 3168 bp in length
 * 134373 134472: gap of 100 bp
 * 134473 136524: contig of 2052 bp in length
 * 136525 137624: gap of 1100 bp
 * 137625 140189: contig of 2565 bp in length
 * 140190 140289: gap of 100 bp
 * 140290 148706: contig of 8417 bp in length
 * 148707 148806: gap of 100 bp
 * 148807 149537: contig of 731 bp in length
 * 149538 169311: gap of 100 bp
 * 169312 169411: contig of 19674 bp in length
 * 169412 219553: contig of 50142 bp in length.
 Location/Qualifiers
 1..219553
 /organism="Homo sapiens"

/mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3p21.3"
 /clone="RP6-153f14"
 ORIGIN
 Query Match 56.3%; Score 704.8; DB 2; Length 219553;
 Best Local Similarity 78.5%; Pred. No. 6.6e-208;
 Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;
 QY 141 TTCAACAATTCAGTGATAACAGCC---AGAGAACAAACGCTTCCTAAAGTCAAGGAG 197
 Db TTAGCAGCTTTCAATGACAGCAGCAGCAGAGAGATCAAGACTTCCCTGAGTTCGCAAG 45907
 QY 198 GTCTTTTGGCCCTGTGTGTACTCTGTAGTGTCTTTTGGACTGCTAGGAAATCCCTG 257
 Db 45908 GTCTTTTGGCCCTGTGTGTACTCTGTAGTGTCTTTTGGTCTGGTGGGAACTCTG 45967
 QY 258 GTTCTGATTAATACATTTTCTACAGAGCTGAGGACTCTGACAGATGTGTTCTGCTG 317
 Db 45968 GTGCTGGTTCATATCCATCTTCTACATAAGTTGAGAGCCCTGACGGATGTGTTCTCGTG 46027
 QY 318 AACTTGGCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCTGCTTTTGGGCTATGACGGC 377
 Db 46028 AACCTACCCCTGGCTGACCTGGTGTGTGTCTGTACTCTGCTGCTTTCTGGGCTATGACGGC 46087
 QY 378 ACCTATGAGTGGGTCTTTGGCAGAGTCAATGTGCAAAACTCTTTGAGGCGATGTATACAATG 437
 Db 46088 ATCCATGAATGGGTGTGTGGCAGGTCAATGTGCAAGAGCCTTACTGGGCATCTACATATT 46147
 QY 438 AACTTCTAGCTGTCCAGTCTCACTCTCACTGTCATACAGTGGATCGTTTCATTTAGTAGT 497
 Db 46148 AACTTCTACAGTCCATGCTCTCTCCTCCTGTCATCACTGTGGAATCGTTTCATTTAGTAGT 46207
 QY 498 GTCCAGGCTACCAAGGCCCTTCAACCGCAGGCTTAAGTGGAGATCTGGGCCAAAGTCATT 557
 Db 46208 GTTAAGGCCCAACAGGCCCTACACAGCAAGCAGGAGGATGACCTGGGGCAAGTCAAC 46267
 QY 558 TGCTTGTCTATTTGGTGGTCTCCTGTTGTTGTTTCTTGGCAGATCACTATGGCCAT 617
 Db 46268 AGCTTGTCTCATCTGGGTGATATCCCTGCTGTTTCTTGGCCCAAAATATCTATGGCAAT 46327
 QY 618 GTTCAAGATATTGACAAGCTTATCTGTCACTACACAGTACAGGAGATATCCACTATGCTT 677
 Db 46328 GTCTTTAATCTGACAAGCTCATATGTGGTTACCATGACAGGCAATTTCCACTGTGCTT 46387
 QY 678 GTTGTATACAGATGACTCTGGGGTCTTCTGCTGCAATGCTCACTATGATCTGTGTCTAC 737
 Db 46388 GTTGCCACCCAGATGACACTGGGGTCTTCTTGTGCACTGCTCACCATGATTTGTCTGTAT 46447
 QY 738 TCAGGCATTATCAAGACCTTCTCTCATGCTCGAAACTTCCAGAGACCAAAATCTCTAAAG 797
 Db 46448 TCAGTCATTAATCAAAACACTGCTTCTGCTGGAGGCTTCCAGAGACACAGATCTCTAAG 46507
 QY 798 ATCATCTTCTTGTAGTGGTGTGTCTGCTGTGACCCAGACACCCCTCAACCTTGGCCATG 857
 Db 46508 ATCATCTTCTTGTGTGTGTCTGCTGTGACCCAGATGCCCTTCAACCTCATGAAG 46567
 QY 858 TTAATCCAAAGTACAGCTGGGAGTCTATACCATATACAGCTTTAAGTATGCCATCGTA 917
 Db 46568 TTCAATCCCGAGACACACTGGGAATACTATGCCATGACAGCTTTTCACTACCATCATG 46627
 QY 918 GTGACAGAGGCTATAGCATACTTTTCGGGCTTGCCTTAAACCTGTACTTTATGCTTTGTT 977
 Db 46628 GTGACAGAGGCCATCGCATACCTGAGGGCCCTGCTTAAACCTGTGCTCTATGCTTTGTC 46687
 QY 978 GGCCTTAAAGTTCGGAGAACGCTCTGGAACCTTATGAGGATATCGCTCTCTCTCAC 1037
 Db 46688 AGCCTGAAGTTTCGAAAGAACTTCTGGAAGAACTTGTGAAGGACATTTGTTGCTTCCCTTAC 46747
 QY 1038 CTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCACAGACTTTGTTGTGCTCC 1097

FEATURES
 source

```

Db      46748  CTTGGGGTCTCACATCAATGTAATCTTCTGAGGACAAATCCACAGACTTTTCTGCTCC 46807
Qy      1098  CACATGTAGAGACACACAGTATGTTCCAAATGTAGTAGGCTTCCACACTTAGAGAAG 1157
Db      46808  CACAATGTGGAGGCCACACAGCATGTTCACAA--GTTATAGGCTTCCACAGGGTTTCGAGAA 46864
Qy      1158  TTAATAACAGAAATTTAGAGCATGGCTGTATCATTTGGATGCAACAAGAAAGCTTTGC 1217
Db      46865  GCTGCTCTGGAATTTGCAAGGCATGGCTGTGCTCTTGAATGTGTGAGGAGGCTTTGT 46924
Qy      1218  TTATAGCATGTGGAGTATCATGGAGAAG 1245
Db      46925  TTATAGCTTGGCATTTCTCATGGAGAAG 46952

RESULT 15
AF124380
LOCUS   Macaca mulatta chemokine receptor bonzo (STRL33) gene, complete
DEFINITION
ACCESSION   AF124380
VERSION     AF124380.1
KEYWORDS    GI:4877754
SOURCE      Macaca mulatta (rhesus monkey)
ORGANISM   Macaca mulatta

REFERENCE
1 (bases 1 to 1032)
AUTHORS    Margulies B.J., Hauer, D.A. and Clements, J.E.
TITLE      Identification and comparison of eleven rhesus macaque chemokine
receptors
JOURNAL    AIDS Res. Hum. Retroviruses 17 (10), 981-986 (2001)
MEDLINE    21354176
PUBMED     11461684
REFERENCE  2 (bases 1 to 1032)
AUTHORS    Margulies, B.J., Hauer, D.A. and Clements, J.E.
TITLE      Direct Submission
JOURNAL    Submitted (27-JAN-1999) Comp. Med., Johns Hopkins University, 720
Rutland Ave., Baltimore, MD 21205, USA

FEATURES
source
1..1032
/organism="Macaca mulatta"
/mo_type="genomic DNA"
/db_xref="taxon:9544"
<1..>1032
/gene="STRL33"
<1..>1032
/gene="STRL33"
/product="chemokine receptor bonzo"
1..1032
/gene="STRL33"
/note="similar to HIV/SIV receptor"
/codon_start=1
/product="chemokine receptor bonzo"
/db_xref="GI:4877755"
/translation="MAEVDHYEDDGLNSQSEEHQDPQLQPRKVFPLPOMLVFV
CGLVNSLVLFISFYHKLSQSDVFLNPLADLVFVCTLPFWAYAGIHEWIFQVM
KTLIVTYINFTSMILTCITVDREIVVVKATKAYNOQAKRWGKVICLLIIVIS
LLSLPQIYGNVFNLDKLCIYHDEISITVLATQMTLGFPLPAMIVCYSIK
LVHAGGFKHRSKLIIFLVMAVFLITOTPFNLKILRSTHWEYVANTSEHYITVEA
IAYLRACLPVLYAFVSLKPRKFNKFLVKDGLCLPLGLVSHQWKSSEDNSKTFASHN
VEATSMFQL"

Query Match 55.6%; Score 696; DB 9; Length 1032;
Best Local Similarity 81.0%; Pred. No. 1.9e-205;
Matches 823; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

Qy      121  ACGAGGGAGATTCTGGCTCTTCAACAATTCAGTGAACAGCC---AGAGAACAAAC 177
Db      17  ACTATGAAGATGATGGGTCTCTCAACAGTTTCAATGACAGCGCAGGAGGAGCATCAAG 76

```

Search completed: September 29, 2004, 02:07:45
Job time : 4916 secs

```

Qy      178  GCTTCTTAAAGTTCAAGAGAGTCTTTTGGCCCTGTGTGTACCTGGTAGTGTGTTGTTGTTG 237
Db      77  ACTTCTTGCAGTTTCAGGAAGTCTTTCTGCCCTGCATGTACCTGGTGGTGTGTTGTCTGTG 136
Qy      238  GACTGCTAGGAATCTCCCTGGTTCTGATTATATACATTTTCTACCAAGAGCTGAGAGCTC 297
Db      137  GCCTGTGGGGAACCTCCCTGGTCTGGTCAATCCATCTTCTACCAATAGTCGAGAGCC 196
Qy      298  TGACAGATGTGTTTCTGCTGAACTTGCCCTCGCTGCTGACCTGGTGTGTTGTTGTTGTTG 357
Db      197  TGACGGACGTGTTCTCTGGTGAACCTACCCCTGGCTGACCTGGTGTGTTGTTGTTGTTG 256
Qy      358  CCTTTTGGGCTATGACAGGCACCTATGAGTGGTCTTTTGGCACAGTCATGTGCAAAACTC 417
Db      257  CCTTCTGGGCTATGAGGCATCCATGAATGGAATCTTTGGCCAGGTCATGTGCAAGACC 316
Qy      418  TTGAGGCATGTATACAATGAATCTTCTAGTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 477
Db      317  TACTGGGCTGCTACACTATTAACCTTCTACACATCCATGCTCATCTCTCACCTGATCACTG 376
Qy      478  TGGATCGTTTCATGTAGTGTCCAGCTACCAAGGCTTCAACGGCAGGCTCAAGTGA 537
Db      377  TGGATCGTTTCATGTAGTGTAAAGCCACCAAGGCTTCAACGGCAGGCTCAAGTGA 436
Qy      538  AGATCTGGGCTCAAGTCAATTTGCTGCTCATTTGGGTGGTCTCTCCCTGTTGGTCTTTTTC 597
Db      437  TGACTTGGGCTCAAGTCAATTTGCTGCTCATCTGGGTGATATCCCTGCTGTTTCTTCTG 496
Qy      598  CACAGATCATCTATGGCATGTTCAAGATATTGACAAAGCTTATCTGTCAAGTACCAAGT 657
Db      497  CCCAAATATATATGGAATGCTTTTAATCTGACAAAGCTCATATGTGTTATCATGACG 556
Qy      658  AGGAGATATCCACTATGTTTGTATACAGATGACTCTGGGGTCTCTCCCTGCCATTG 717
Db      557  AGGAGATTTCCACTGTGTTCTTGCCACCCAGATGACACTGGGGTCTCTTCTGCTGCTG 616
Qy      718  TCACTATGATCTGCTACTCAGGATTAATCAAGACCTTGTCTATGTTCGAAATTC 777
Db      617  TCGCCATGATGTTCTGCTATTCAGTCAATAATCAAAACACTGCTTTCATGTGGAGGCTTCC 676
Qy      778  AGAAGCACAAATCTCTAAAGATCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 837
Db      677  AGAAGCACAAATCTCTAAAGATCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 736
Qy      838  CACCTTTCAACCTTGGCATGTTAATCCAAAGTACAAAGCTGGGAGTACTATACCAATACCA 897
Db      737  CACCTTTCAACCTTGGCATGTTAATCCGAGCACACACTGGGAGTACTATGCCATGACCA 796
Qy      898  GCTTTAAGTATGCCATCGTAGTGACAGAGCTATAGCATCTTTCGGGCTTGCCTTAACC 957
Db      797  GCTTTCACTACACCATCATAGTGACAGAGGCCATCGCATACCTGAGGGCTGCTGCTTAACC 856
Qy      958  CTGTACTTTATGCTTTTGTGGCTTAAAGTTCGGGAAGAACGCTCTGGAACCTTATGAAG 1017
Db      857  CTGTGCTCTATGCTTTGTGACGCTGAAGTTTCGAAGAACCTTCTGGAACCTTGTGAAG 916
Qy      1018  ATATCGGCTGCCCTCTCTACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTT 1077
Db      917  ACATGGCTGTCTCTCTTACCTTGGGCTTCTCATCAATGGAATCTTCTGAGGACAAAT 976
Qy      1078  CCAAGACTTGTCTGCTCTCCCAATAGTAGAGACCAAGTATGTTCCCAATTGTAG 1133
Db      977  CCAAGACTTGTCTGCTCTCCCAATAGTAGAGGCCACCAAGCATGTTTCCAGCTATAG 1032

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: September 29, 2004, 00:26:24 / Search time 522 Seconds
(without alignments)
10181.026 Million cell updates/sec

Title: US-10-603-141-1
Perfect score: 1251
Sequence: 1 gaattcgacagggcaga.....gtatcatggagaagcctgac 1251

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 337863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1210	96.7	1475	2	AAV43795	Aav43795 Rodent ch
2	1053	84.2	1053	2	AX86033	AX86033 Membrane
3	1053	84.2	1053	2	AX86032	AX86032 Nucleic a
4	706.4	56.5	2238	2	AAV24017	Aav24017 Human HBM
5	705.4	56.4	1953	7	ABZ42779	ABZ42779 Human SIV
6	704.8	55.3	1918	2	AAV68515	AAV68515 STRL33 co
7	689.4	55.1	1029	2	AAx07291	Aax07291 Pig-taile
8	689.4	55.1	1029	5	AAx07291	Aax07291 Pig-taile
9	689.4	55.1	1029	6	ABN84841	Abn84841 Pig-taile
10	685	54.8	1037	2	AAx07290	Aax07290 African g
11	685	54.8	1037	5	AAx07615	Aax07615 Green mon
12	685	54.8	1037	6	ABN84640	Abn84640 African g
13	680.8	54.4	1029	6	ABN84640	Abn84640 African g
14	679.2	54.3	1029	4	AAx07339	Aax07339 CDNA enco
15	679.2	54.3	1029	9	AAx07339	Aax07339 CDNA enco
16	679.2	54.3	1029	5	AAH77725	Aah77725 Nucleotid
17	676.8	54.1	1026	2	AAZ28362	Aaz28362 Platelet
18	676	54.0	1029	6	ABN84640	Abn84640 African g
19	648	51.8	1026	2	AAx07289	Aax07289 Human HIV
20	648	51.8	1026	5	AAx07614	Aax07614 Human CDN
21	648	51.8	1026	6	ABN84639	Abn84639 Human G-p
22	471.8	37.7	581	2	AAV43794	Aav43794 Rodent ch
23	183.4	14.7	1074	3	AAA91745	Aaa91745 Human G p

24	183.4	14.7	1074	9	ADC22542	Adc22542 Human G p
25	183.4	14.7	2462	7	ABZ42725	Abz42725 Human C-C
26	183.4	14.7	2577	3	AAA88560	Aaa88560 Human CC
27	183.4	14.7	2577	6	AAD32916	Aad32916 Homo sapi
28	183.4	14.7	2577	7	ACA56497	Acas6497 Human sig
29	183.4	14.7	2577	7	ABX15306	Abx15306 Human CC
30	181.2	14.5	2751	2	AAQ66164	Aaq66164 Seven tra
31	181.2	14.5	2751	3	AAV18351	Aavi18351 Murine V3
32	181.2	14.5	2751	3	AAA91711	Aaa91711 Murine 7T
33	181.2	14.5	2751	6	ABX54240	Abx54240 Mouse 7 t
34	180.2	14.4	1110	5	AB197980	Abi97980 Non-endog
35	178.6	14.3	1074	3	AAA30715	Aaa30715 DNA encod
36	178.6	14.3	1074	9	ADC22694	Adc22694 Human G p
37	177.8	14.2	1137	7	AAU56092	Aau56092 Mouse cmk
38	177.8	14.2	1137	8	ADA02454	Ada02454 Mouse cmk
39	177.8	14.2	1137	9	ADB72193	Adb72193 Mouse cmk
40	177.8	14.2	2072	6	AA317156	Aaa17156 Mouse lym
41	177.8	14.2	2072	7	AAU56091	Aau56091 Mouse cmk
42	177.8	14.2	2072	8	ADA02453	Ada02453 Mouse cmk
43	177.8	14.2	2072	9	ADB72192	Adb72192 Mouse cmk
44	177.8	14.2	2084	6	ABK54274	Abk54274 Mouse 7 t
45	177.8	14.2	2085	3	AAA91745	Aaa91745 7TM recep

ALIGNMENTS

RESULT 1
AAV43795
ID AAV43795 standard; cDNA; 1475 BP.

AC AAV43795;

XX 20-OCT-1998 (first entry)

XX Rodent chemokine receptor 941D12 nucleotide sequence.

Chemokine; primate; human; rodent; chemokine receptor; asthma;
inflammatory response; immune response; leukocyte migration; GPCR;
leukocyte adhesion; chemottractant; modulation; antiviral response;
cellular morphology modification response; G-protein coupled receptor;
phosphoinositide lipid turnover; abnormal proliferation; regeneration;
atrophy; 941D12; ss.

OS Mus sp.

XX Key Location/Qualifiers
FT CDS 1..975

FT FT /tag= a /product= "chemokine receptor 941D12"

FT misc_feature 942

FT FT /tag= b /note= "nucleotide designated C may be C or T"

FT FT /tag= c /note= "1412..1422"

FT FT /tag= c /note= "nucleotides 1412 and 1422 each designated c may be A, C, G or T"

XX WO9832858-A2.

XX 30-JUL-1998.

XX 22-JAN-1998; 98WO-US000902.

XX 23-JAN-1997; 97US-0036715P.

XX (SCHE) SCHERING CORP.

XX PI Mattson JD, Soto-Trejo H, Hedrick JA, Gorman DM, Zlotnik A;

XX WPI; 1998-427954/36.

XX P-PSDB; AAW70001.

PT Rodent and primate chemokines and chemokine receptors - useful
 PT diagnostically and therapeutically to treat conditions associated with
 XX abnormal physiology or development e.g. inflammatory conditions.
 XX
 PS Example 3; Page 97-99; 105pp; English.

XX This represents a rodent chemokine receptor 941D12 nucleotide sequence.
 CC The invention provides novel primate and rodent chemokines and chemokine
 CC receptors. The chemokines, receptors and binding compounds (optionally
 CC antibodies/fragments specifically binding the chemokines) are useful
 CC therapeutically to treat conditions associated with abnormal physiology
 CC or development e.g. inflammatory conditions such as asthma. Chemokines
 CC are important in immune and inflammatory responses in that they induce
 CC leukocyte migration and adhesion. They are also chemoattractants for
 CC several cells involved in inflammation and can induce other biological
 CC responses e.g. modulation of second messenger levels (e.g. Ca++),
 CC cellular morphology modification responses, phosphoinositide lipid
 CC turnover, possible antiviral responses etc. The chemokine receptors of
 CC the invention exhibit structural properties of G-protein coupled
 CC receptors (GPCR), although their ligands have not yet been identified.
 CC The chemokine and chemokine receptor polypeptides are useful to produce
 CC ligand:receptor complexes in vivo or in assay techniques. Assays may also
 CC involve chemical antagonists which block complex production or utilize
 CC competitive binding. Binding compounds identified (agonists or
 CC antagonists) can be used to modulate the physiological responses in cells
 CC (especially neurons, macrophages or lymphocytes) to treat e.g. abnormal
 CC proliferation, regeneration, generation and atrophy. The polypeptides are
 CC also used to produce antibodies useful diagnostically, for drug screening
 CC or for polypeptide purification. The polynucleotides are useful to
 CC produce probes for detecting the polypeptides, and to isolate the
 CC polypeptides or related sequences, especially from other species. They
 CC also allow transformation of cells for polypeptide production
 XX

SQ Sequence 1475 BP; 374 A; 346 C; 312 G; 443 T; 0 U; 0 Other;

Query Match 96.7%; Score 1210; DB 2; Length 1475;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1225; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 QY 15 GGCAGACCTTTAGTGACACACTTCACTCGGAACAAAGCTACTGGGCTTCTCTCTGAT 74
 DB 4 GGCAGACCTTTAGTGACACACTTCACTCGGAACAAAGCTACTGGGCTTCTCTCTGAT 63
 QY 75 GCCATGATGATGGGCATCAAGAGTCAGCTCTGTACGATGGGCACCTACGAGGAGATTTC 134
 DB 64 GCCATGATGATGGGCATCAAGAGTCAGCTCTGTACGATGGGCACCTACGAGGAGATTTC 123
 QY 135 TGGCTCTTCAACAATCCAGTGATACAGCCAGGAGACAAACGGCTTCCTAAAGTTCAAG 194
 DB 124 TGGCTCTTCAACAATCCAGTGATACAGCCAGGAGACAAACGGCTTCCTAAAGTTCAAG 183
 QY 195 GAGGCTCTTTGGCCCTGTGTGACCTGGTAGTGTGTCTTGGACTGTAGGAACCTCC 254
 DB 184 GAGGCTCTTTGGCCCTGTGTGACCTGGTAGTGTGTCTTGGACTGTAGGAACCTCC 243
 QY 255 CTGGTCTGATATATACATTTCTACCAAGAGCTGAGGACTCTGACAGATGTTCTG 314
 DB 244 CTGGTCTGATATATACATTTCTACCAAGAGCTGAGGACTCTGACAGATGTTCTG 303
 QY 315 CTGAACCTTGCCCTGGCTGACCTGGTGTGTCTGTACTCTGCTCTTGGGCTTATGCA 374
 DB 304 CTGAACCTTGCCCTGGCTGACCTGGTGTGTCTGTACTCTGCTCTTGGGCTTATGCA 363
 QY 375 GGCACCTATGAGTGGGTCTTTGGCAAGTCATGTGCAAAACCTCTTCGAGGATGATACA 434
 DB 364 GGCACCTATGAGTGGGTCTTTGGCAAGTCATGTGCAAAACCTCTTCGAGGATGATACA 423
 QY 435 ATGAACCTTCTACGTGTCCATGCTCACTCTCACTGCATCACAGTGCATGTTTCAATTGTA 494
 DB 424 ATGAACCTTCTACGTGTCCATGCTCACTCTCACTGCATCACAGTGCATGTTTCAATTGTA 483
 QY 495 GTGGTCCAGGCTACCAAGGCTTCAACCGGAGGCTTAAGTGAAGATCTGGGCGCAAGTC 554

DB 484 GTGGTCCAGGCTACCAAGGCTTCAACCGGAGGCTAAGTGGAAAGATCTGGGCGCAAGTC 543
 QY 555 ATTTGCTTGTCTCATTTGGGTGTCTCCCTGTGTGGTTTCTTTGGCCACAGATCATCTATGGC 614
 DB 544 ATTTGCTTGTCTCATTTGGGTGTCTCCCTGTGTGGTTTCTTTGGCCACAGATCATCTATGGC 603
 QY 615 CATGTTCAAGATATTGACAAGCTTATCTGTCAAGTACCAAGTACAGTACAGGAGATATCCACTATG 674
 DB 604 CATGTTCAAGATATTGACAAGCTTATCTGTCAAGTACCAAGTACAGTACAGGAGATATCCACTATG 663
 QY 675 GTTCTGTGTTACAGATCACTCTGGGTCTTCTCTGCAATGCTCACTATGATCTGTGC 734
 DB 664 GTTCTGTGTTACAGATCACTCTGGGTCTTCTCTGCAATGCTCACTATGATCTGTGC 723
 QY 735 TACTCAGGATATCAAGACCTTCTCTCATGCTCGAAACTTCCAGAGCAACAAATCTCTA 794
 DB 724 TACTCAGGATATCAAGACCTTCTCTCATGCTCGAAACTTCCAGAGCAACAAATCTCTA 783
 QY 795 AAGATCATCTTCTTGTAGTGGCTGTGCTGCTGACCCAGACACCCCTTCAACCTTGGCC 854
 DB 784 AAGATCATCTTCTTGTAGTGGCTGTGCTGCTGACCCAGACACCCCTTCAACCTTGGCC 843
 QY 855 ATGTTAATCCAAAGTACAAAGCTGGGAGTACTATACCAATACCAAGCTTTAAGTATGCCATC 914
 DB 844 ATGTTAATCCAAAGTACAAAGCTGGGAGTACTATACCAATACCAAGCTTTAAGTATGCCATC 903
 QY 915 GTAGTGACAGAGGCTATAGCATCTTTC - GGGCTTGCCTTAACCCGTACTTTATGCCCT 972
 DB 904 GTAGTGACAGAGGCTATAGCATCTTTC - GGGCTTGCCTTAAACCCGTACTTTATGCCCT 963
 QY 973 TTGTTGGCTTAAAGTTCGGAAGAACCTCTGGAACCTTATGAAGGATATCGGCTGCTCT 1032
 DB 964 TTGTTGGCTTAAAGTTCGGAAGAACCTCTGGAACCTTATGAAGGATATCGGCTGCTCT 1023
 QY 1033 CTCACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTAGGACAGTTCGAGGACTTTGTTCTG 1092
 DB 1024 CTCACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTAGGACAGTTCGAGGACTTTGTTCTG 1083
 QY 1093 CCTCCCAATGTAGACACCAAGTATGTTCCAAATGTTAGTGGCTTCCGACACTTAG 1152
 DB 1084 CCTCCCAATGTAGACACCAAGTATGTTCCAAATGTTAGTGGCTTCCGACACTTAG 1143
 QY 1153 AGAAGTTAATAACAGAAATCTAGGAGCATGGCTGTATCATTTGGATGCAACAAAGAAAGC 1212
 DB 1144 AGAAGTTAATAACAGAAATCTAGGAGCATGGCTGTATCATTTGGATGCAACAAAGAAAGC 1203
 QY 1213 TTTGCTTATAGCATGTGGAGTATCATGGAGAA 1244
 DB 1204 TTTGCTTATAGCATGTGGAGTATCATGGAGAA 1235

RESULT 2

AAx86033/c
 ID AAx86033 standard; cDNA to mRNA; 1053 BP.

XX AAx86033;

DT 15-SEP-1999 (first entry)

XX Membrane penetrating type receptor protein Et60 nucleic acid fragment.

XX 7 times membrane-penetrating type receptor protein; Et60;

XX leukocyte function; control; drug; ss.

XX Unidentified.

OS JP1155573-A.

PN 15-JUN-1999.

XX 27-NOV-1997; 97JP-00325823.

XX 27-NOV-1997; 97JP-00325823.

XX PA (ASAH) ASahi KASEI KOGYO KK.
XX DR WPI; 1999-398070/34.
XX PT New 7 times membrane-penetrating type receptor protein ET60 - useful for
XX PT detecting drug controlling function of leukocyte.
XX PS Claim 6; Page 19; 20pp; Japanese.
XX CC The present sequence represents a fragment of nucleic acid encoding a 7
XX CC times membrane-penetrating type receptor protein ET60. The 7 times
XX CC membrane-penetrating type receptor protein ET60 can be used for detecting
XX CC a drug controlling the function of leukocyte.
XX SQ Sequence 1053 BP; 317 A; 239 C; 253 G; 244 T; 0 U; 0 Other;
Query Match 84.2%; Score 1053; DB 2; Length 1053;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 78 ATGGATGATGGGATCAAGAGTCAAGTCTGTACGATGGGCACTACGAGGAGATTTCTGG 137
Db 1053 ATGGATGATGGGATCAAGAGTCAAGTCTGTACGATGGGCACTACGAGGAGATTTCTGG 994
QY 138 CTCTTCAACAATTCAGTGTAAACAGCCAGAGAACAAACCGTTCCTAAAGTTCAGGAG 197
Db 993 CTCTTCAACAATTCAGTGTAAACAGCCAGGAGAACAAACCGTTCCTAAAGTTCAGGAG 934
QY 198 GTCTTTTGGCCCTGTGCTACTGCTGTAGTGTCTCTTTGACCTGTAGGAATCTCCCTG 257
Db 933 GTCTTTTGGCCCTGTGCTACTGCTGTAGTGTCTCTTTGACCTGTAGGAATCTCCCTG 874
QY 258 GTTCTGATTATATACATTTTCTACAGAAAGTGGAGCTCTGACAGATGTCTTTCTGCTG 317
Db 873 GTTCTGATTATATACATTTTCTACAGAAAGTGGAGCTCTGACAGATGTCTTTCTGCTG 814
QY 318 AACTTGGCCCTGGCTGACTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 377
Db 813 AACTTGGCCCTGGCTGACTGTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 754
QY 378 ACCTATGATGGTCTTTGGGACAGTCACTGTGCAAAACTTTCGAGGATGTATACAAATG 437
Db 753 ACCTATGATGGTCTTTGGGACAGTCACTGTGCAAAACTTTCGAGGATGTATACAAATG 694
QY 438 AACTTCTACGTGTCCATGCTCACTCTACCTGATCACTACAGTGAATCGTTTCATTTAGTG 497
Db 693 AACTTCTACGTGTCCATGCTCACTCTACCTGATCACTACAGTGAATCGTTTCATTTAGTG 634
QY 498 GTCCAGGCTACCAAGCCCTCAACCGGAGGCTAAGTGAAGATCTGGGGCCAAAGTCAAT 557
Db 633 GTCCAGGCTACCAAGCCCTCAACCGGAGGCTAAGTGAAGATCTGGGGCCAAAGTCAAT 574
QY 558 TGCTTGTCTCATTTGGGTGTCTCCCTGTGTGTGTCTTTTCCACAGATCATCTATGGCCAT 617
Db 573 TGCTTGTCTCATTTGGGTGTCTCCCTGTGTGTGTCTTTTCCACAGATCATCTATGGCCAT 514
QY 618 GTTCAAGATATTGACAGCTTATCTGTGACGACCACTGAGGAGATATCCACTATGTT 677
Db 513 GTTCAAGATATTGACAGCTTATCTGTGACGACCACTGAGGAGATATCCACTATGTT 454
QY 678 CTGTGTTATACAGATGACTCTGGGGTCTTCTCCGCAATGCTCACTATGATTTGTGCTAC 737
Db 453 CTGTGTTATACAGATGACTCTGGGGTCTTCTCCGCAATGCTCACTATGATTTGTGCTAC 394
QY 738 TCAGGCAATTAACAAGACCTTGTATGCTGAAACTTCCAGAGACAAATCTCTTAAG 797
Db 393 TCAGGCAATTAACAAGACCTTGTATGCTGAAACTTCCAGAGACAAATCTCTTAAG 334
QY 798 ATCATCTTCTGTAGTGGCTGTGTCTGCTGACCCAGACACCTTCAACCTTCCCATG 857
Db 333 ATCATCTTCTGTAGTGGCTGTGTCTGCTGACCCAGACACCTTCAACCTTCCCATG 274

QY 858 TTAATCCAAAGTACAAGCTGGGAGTACTATACCAATACAGAGCTTTAAGTATGCCATCGTA 917
Db 273 TTAATCCAAAGTACAAGCTGGGAGTACTATACCAATACAGAGCTTTAAGTATGCCATCGTA 214
QY 918 GTGACAGAGGCTATAGCATACTTTCGGGCTTCGCTTAACCCCTGTACTTTATGCCCTTTGTT 977
Db 213 GTGACAGAGGCTATAGCATACTTTCGGGCTTCGCTTAACCCCTGTACTTTATGCCCTTTGTT 154
QY 978 GGTCTTAAAGTTCGGGAAGAACCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCTCAC 1037
Db 153 GGTCTTAAAGTTCGGGAAGAACCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCTCAC 94
QY 1038 CTGGAGTCTCAAGTCAATGAAGTCTTCTGAGGACAGTTCACAGACTTGTTCGCTCC 1097
Db 93 CTGGAGTCTCAAGTCAATGAAGTCTTCTGAGGACAGTTCACAGACTTGTTCGCTCC 34
QY 1098 CACAATGTAGAGACCAACCACTATGTTCCAATTG 1130
Db 33 CACAATGTAGAGACCAACCACTATGTTCCAATTG 1
RESULT 3
AAx86032
ID AAx86032 standard; cDNA to mRNA; 1053 BP.
XX AC AAx86032;
XX DT 15-SEP-1999 (first entry)
XX DE Nucleic acid encoding a membrane penetrating type receptor protein ET60.
XX KW 7 times membrane-penetrating type receptor protein; ET60;
XX KM leukocyte function; control; drug; ss.
XX OS Unidentified.
XX PN JP11155573-A.
XX PD 15-JUN-1999.
XX PF 27-NOV-1997; 97JP-00325823.
XX PR 27-NOV-1997; 97JP-00325823.
XX PA (ASAH) ASahi KASEI KOGYO KK.
XX DR WPI; 1999-398070/34.
XX DR P-PSDB; AAY23825.
XX PT New 7 times membrane-penetrating type receptor protein ET60 - useful for
XX PT detecting drug controlling function of leukocyte.
XX PS Claim 4; Page 18-19; 20pp; Japanese.
XX CC The present sequence encodes a 7 times membrane-penetrating type receptor
XX CC protein ET60. The 7 times membrane-penetrating type receptor protein ET60
XX CC can be used for detecting a drug controlling the function of leukocyte
XX SQ Sequence 1053 BP; 244 A; 253 C; 239 G; 317 T; 0 U; 0 Other;
Query Match 84.2%; Score 1053; DB 2; Length 1053;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1053; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 78 ATGGATGATGGGATCAAGAGTCAAGTCTGTACGATGGGCACTACGAGGAGATTTCTGG 137
Db 1 ATGGATGATGGGATCAAGAGTCAAGTCTGTACGATGGGCACTACGAGGAGATTTCTGG 60
QY 138 CTCTTCAACAATTCAGTGTAAACAGCCAGGAGAACAAACCGTTCCTAAAGTTCAGGAG 197
Db 61 CTCTTCAACAATTCAGTGTAAACAGCCAGGAGAACAAACCGTTCCTAAAGTTCAGGAG 120
QY 198 GTCTTTTGGCCCTGTGCTACTGCTGTAGTGTGTCTTTTGGACTGTAGGAACTCCCTG 257

```
Db 121 GTCCTTTGGCCCTGTGTACTCTGGTAGTGTGTTCTTTGGAGCTGTAGAAATCCCTG 180
Qy 258 GTTCTGATTATATACATTTTCTACAGAGCTGAGGACTCTGCAGAGTGTGTTCTGCTG 317
Db 181 GTTCTGATTATATACATTTTCTACAGAGCTGAGGACTCTGCAGAGTGTGTTCTGCTG 240
Qy 318 AACTTGCCCTGGCTGAGCTGGTGTGTTGTCGTACTCTGCGCTTTTGGGCTATGAGGC 377
Db 241 AACTTGCCCTGGCTGAGCTGGTGTGTTGTCGTACTCTGCGCTTTTGGGCTATGAGGC 300
Qy 378 ACCTATGAGTGGGTCTTTGGCACAGTCAATGTGCAAACTCTCGAGGCATGTATACAATG 437
Db 301 ACCTATGAGTGGGTCTTTGGCACAGTCAATGTGCAAACTCTTCAGGCAATGTATACAATG 360
Qy 438 AACTTCTAGTGTCCATGCTCACTCTCACCTGCAATCACAGTGGATCGTTTCATTTAGTG 497
Db 361 AACTTCTAGTGTCCATGCTCACTCTCACCTGCAATCACAGTGGATCGTTTCATTTAGTG 420
Qy 498 GTCCAGGCTACCAAGGCTTTCACCCGCGAGCTAAGTGGAGATCTGGGCGCAAGTCATT 557
Db 421 GTCCAGGCTACCAAGGCTTTCACCCGCGAGCTAAGTGGAGATCTGGGCGCAAGTCATT 480
Qy 558 TGCTTGCTCATTTGGGGTGTCTCCCTGTTGGTTTCTTTGGCACAGATCATCTATGCCAT 617
Db 481 TGCTTGCTCATTTGGGGTGTCTCCCTGTTGGTTTCTTTGGCACAGATCATCTATGCCAT 540
Qy 618 GTTCAAGATATTGACAGCTTATCTGTCACTACACAGTGGAGATATCCATATGGTT 677
Db 541 GTTCAAGATATTGACAGCTTATCTGTCACTACACAGTGGAGATATCCATATGGTT 600
Qy 678 CTTGTTATACAGATGACTCTGGGGTCTTCTGCGCAATGCTCACTATGATTTGCTCTAC 737
Db 601 CTTGTTATACAGATGACTCTGGGGTCTTCTGCGCAATGCTCACTATGATTTGCTCTAC 660
Qy 738 TCAGGCAATTATCAAGACTTCTTCACTGCTCGAAACTTCCAGAAAGCACAAATCTTAAAG 797
Db 661 TCAGGCAATTATCAAGACTTCTTCACTGCTCGAAACTTCCAGAAAGCACAAATCTTAAAG 720
Qy 798 ATCATCTCTCTGTAGTGGCTGTGTTCTCTGCTCACCCAGACACCCCTTCAACCTTGCATG 857
Db 721 ATCATCTCTCTGTAGTGGCTGTGTTCTCTGCTCACCCAGACACCCCTTCAACCTTGCATG 780
Qy 858 TTAATCCAAAGTACAAGCTGGGAGTACTATACATTAACAGCTTTAAGTATGCCATCGTA 917
Db 781 TTAATCCAAAGTACAAGCTGGGAGTACTATACATTAACAGCTTTAAGTATGCCATCGTA 840
Qy 918 GTCCAGAGGCTATAGCATACTTTCGGGCTTGGCTTAAACCTGTACTTTATGCTTGGT 977
Db 841 GTCCAGAGGCTATAGCATACTTTCGGGCTTGGCTTAAACCTGTACTTTATGCTTGGT 900
Qy 978 GGCTTAAAGTTCGGGAGAACGCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCTAC 1037
Db 901 GGCTTAAAGTTCGGGAGAACGCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCTAC 960
Qy 1038 CTGGAGTCTCAAGTCAATGGAGTCTTCTGAGGACAGTTCAGAGCTTGTCTGCTCTCC 1097
Db 961 CTGGAGTCTCAAGTCAATGGAGTCTTCTGAGGACAGTTCAGAGCTTGTCTGCTCTCC 1020
Qy 1098 CACAATGTAGAGACCAACCACTATGTTTCCAATTG 1130
Db 1021 CACAATGTAGAGACCAACCACTATGTTTCCAATTG 1053
```

RESULT 4

AAV24017

ID AAV24017 standard; cDNA; 2238 BP.

XX

AC AAV24017;

XX

DT 07-AUG-1998 (first entry)

XX

DE Human HMBU14 coding sequence.

```
XX KW Human; 7-transmembrane receptor; HMBU14; infection; HIV; cancer; asthma;
KW anorexia; Parkinson's disease; acute heart failure; atherosclerosis;
KW osteoporosis; angina pectoris; myocardial infarction; psychotic disorder;
KW benign prostatic hypertrophy; neurological disorder; therapy;
KW Huntington's disease; Gilles de la Tourette's syndrome; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 332..1360
FT /*tag= a
FT /product= "HMBU14"
XX PN EP834563-A2.
XX PD 08-APR-1998.
XX PF 23-SEP-1997; 97EP-00307428.
XX PR 26-SEP-1996; 96US-0026669P.
XX PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX PI Elshourbagy NA, Bergsma DJ, Ellis CE;
XX DR WPI: 1998-208924/19.
XX DR P-PSDB; AAW54041.
XX PT New isolated human 7-trans-membrane receptor, HMBU14 - useful for
PT developing products for treating, e.g. infections, pain, cancers,
PT anorexia, hypotension, osteoporosis and asthma.
XX PS Claim 4; Fig 1; 36pp; English.
```

```
CC This sequence encodes the human HMBU14 protein of the invention. HMBU14
CC is a human 7-transmembrane (TM) receptor. The products can be used for
CC treating or preventing conditions related to abnormal HMBU14 expression
CC or activity, e.g. infections such as bacterial, fungal, protozoan and
CC viral infections especially infections caused by HIV-1 and HIV-2, pain,
CC cancers, anorexia, bulimia, asthma, Parkinson's disease, acute heart
CC failure, atherosclerosis, hypotension, hypertension, urinary retention,
CC osteoporosis, angina pectoris, myocardial infarction, ulcers, asthma,
CC allergies, benign prostatic hypertrophy and psychotic and neurological
CC disorders, including anxiety, schizophrenia, manic depression, delirium,
CC Huntington's disease or mental retardation, and dyskinesias, such as
CC dementia.
CC Huntington's disease or Gilles de la Tourette's syndrome. The products
CC can also be used for diagnosing a disease or susceptibility to a disease
CC related to the expression of HMBU14 and for identifying compounds which
CC bind to and activate or inhibit a receptor for the polypeptide
```

XX SQ Sequence 2238 BP; 593 A; 512 C; 517 G; 616 T; 0 U; 0 Other;

Query Match 56.58; Score 706.4; DB 2; Length 2238;
Best Local Similarity 78.6%; Pred. No. 3.2e-209;
Matches 871; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

```
Qy 141 TTCAACAAATTCAGTGATAACAGCC---AGGAGAACAAACGCTTCTCTAAAGTTCAAGGAG 197
Db 365 TTCAAGCATTTCAATGACAGCAGCCAGGAGGAGCATCAAGACTTCTGCGAGTTCAGCAAG 424
Qy 198 GTCTTTTGGCCCTGTGTACCTGGTAGTGTGTTGCTTTGGACTGTAGAAATCTCCTG 257
Db 425 GTCTTTTGGCCCTGTGTACCTGGTAGTGTGTTGCTTTGGACTGTAGAAATCTCCTG 484
Qy 258 GTTCTGATTATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTGTTCTGCTG 317
Db 485 GTGCTGTGCATATCCATCTTCTACCATAGTTCAGAGCTGACGATGTGTTCTGCTG 544
Qy 318 AACTTGCCCTGGCTGACCTGGTGTGTTGCTGTACTCTGCTGCTTTTGGGCTATGAGGC 377
Db 545 AACCTACCCCTGGCTGACCTGGTGTGTTGCTGTCACTCTGCTCTTCTGGGCTATGAGGC 604
```


Qy	258	GTTCTGATTATATACATTTTCTACAGAAAGCTGAGGACTCTGACAGATGTGTTTCTGCTG	317
Db	235	GTGCTGGTCATATCCATCTTCTACCATAAAGTTGAGAGCCCTGACGGATGTGTTCTCTGGT	294
Qy	318	AACTTGCCCCCTGGGTGACCTGGTGTGTTGTCGTACTCTGCCCCTTTTGGCCCTATGCAAGC	377
Db	295	AACCTAACCCCTGGGTGACCTGGTGTGTTGTCGTACTCTGCCCCTTCTGGCCCTATGCAAGC	354
Qy	378	ACCTATGAGTGGGTCTTTGGCACAGTCAATGTCGAAACTCTTCAGAGCATGTATACAAATG	437
Db	355	ATCCATGAATGGGTGTTTGGCCAGGTCAATGTCGAAGCCCTACTGGGCATCTACACATATT	414
Qy	438	AACTTCTACGTGTCATGCTCACTCTCACCTGTCATCAAGTGGATCGTTTTCAATTGTAGTG	497
Db	415	AACTTCTACAGTCCATGCTCATCTCACTGCATCACTGTGGATCGTTTTCAATTGTAGTG	474
Qy	498	GTCAGAGCTACCAAGGCTTCAACGGCAGAGCTAAGTGGAGATCTCGGGCCAACTCATTT	551
Db	475	GTTAAGGCCACCAAGGCCCTACACAGCAGCCGAAGAGGATGACCTGGGGCAAGGTCAACC	534
Qy	558	TGCTTGCTCATTTGGTGGTCTCCCTGTTGTTTCTTTGGCACAGATCATCTATGGCCAT	617
Db	535	AGCTTGCTCATCTGGGTGATATCCCTGCTGGTTTCTTGCCCCAAATATCTATGGCAAT	594
Qy	618	GTTCAAGATATTGACAAGCTTATCTGTCAGTACCACAGTAGGAGAGATATCCACTATCGTT	677
Db	595	GTCTTTAATCTGCACAGCTCATATGTGGTTACCATGACGAGGCAATTTCCACTGTGGTT	654
Qy	678	CTTGTTATACAGATGACTCTGGGGTCTTCTGTGCCATGTCTCACTATGATTTCTGTGTCTAC	737
Db	655	CTTGCCACCCAGATGACACTGGGGTCTTCTTGCCACTGTCTCACTGATTTGTCTGTCTAT	714
Qy	738	TCAGGCATATCAAGACCTTGCTTCATGCTCGAAACTTCCAGAGACACAATCTCTAAAG	797
Db	715	TCAGTCATATCAAAACACTGCTTTCATGCTGGAGGCTTCCAGANGCACAGATCTCTAAG	774
Qy	798	ATCATCTTCTTGTTAGTGGCTGTGTTTCTGTGTGACCCAGACACCCTTCAACCTTGCCCATG	857
Db	775	ATCATCTTCTGTGTGATGGCTGTGTTTCTGTGTACCCAGATGCCCCTTCAACCTCATGAAG	834
Qy	858	TTAATCCAAAGTACAAGCTGGAGTACTATACCATAACCAAGCTTTAAGTATGCCATCGTA	917
Db	835	TTCATCCGAGCACACACTGGGAATACTATGCCATGACCAAGCTTCTACTACCACTCATG	894
Qy	918	GTGACAGAGGTATAGCATACTTTGCGGCTTGCCTTAAACCCTGTACTTTATGCTTTGTT	977
Db	895	GTGACAGAGCCATCGCATACCTGAGGGCTGCGCTTAAACCCTGTGCTATGCTCTTGTC	954
Qy	978	GGCTTAAAGTTCCGAAAGAACGTCGAAACTTATGAAGGATATCGGCTGCCTCTCTCAC	1037
Db	955	AGCCTGAAGTTTCGAAAGAACTTCTGAAACTTGTGAAGGACATGTTGTTGCTCCCTTAC	1014
Qy	1038	CTGGAGTCTCAAGTCAATGGAGTCTTCTGAGGACAGTTCAGAGACTTGTCTGCTGCC	1097
Db	1015	CTTGCGGTCTCAACATCAATGGAATCTCTGAGGACAAATCCGAAGACTTTTCTGCTGCC	1074
Qy	1098	CACAAATGAGAGACCACAGTATGTTCCAAATTTGATGAGCCCTTGCCACACTTAGAGAAG	1157
Db	1075	CACAAATGAGAGCCACAGCATGTTCCA--GTTATAGGCCTTGCCAGGGTTTCGAGAA	1131
Qy	1158	TTAATAACAAATCTAGGAGCATGGCTGTATCATTTGGATGCAACAGAAAGACTTTGC	1217
Db	1132	CGTGCTCTGGAATTTGCAAGTCAATGGCTGTGCCCTCTTGAATGTGGTGAAGCAGCTTTGT	1191
Qy	1218	TTATAGCATGTGAGTATCATGGAGAAG	1245
Db	1192	TTATAGCTTGGCATTTCTATGAGAG	1219

RESULT 6
AAV68515
ID AAV

[illegible]

```
Db 364 AACTTCTACAGTCCTCATCTCCTCACCTGCAATCACTGTGGATCGTTTCAITGTAGTG 423
Qy 498 GTCAGGCTACCAAGGCTTCAACCCGAGGCTTAAGTGAAGATCTGGGGCCAAAGTATT 557
Db 424 GTTAAGGCCACCAAGGCTTCAACCCGAGGCTTAAGTGAAGATCTGGGGCCAAAGTATT 483
Qy 558 TGGTGTCTCATTTGGGTGGTCTCCCTGTGTGGTCTTCTTGGCCACAGATCATCTATGGCCAT 617
Db 484 AGCTTGTCTCATTTGGGTGGTCTCCCTGTGTGGTCTTCTTGGCCACAAATTAATATGGCAAT 543
Qy 618 GTTCAAGATATTGAACAAGCTTATCTGCAGTACACAGATGAGGAGATATCCACTATGGTT 677
Db 544 GTCTTTAATCTCGAACAAGCTTATCTGCAGTACACAGATGAGGAGATATCCACTATGGTT 603
Qy 678 CTTGTTATACAGATCTCGGCTTCTTCTGCAATGCTCACTATGATTTCTGTGTAC 737
Db 604 CTTGCCACCCAGATGACACTGGGCTTCTTCTGCACTGCTCACCATGATTTCTGTGTAT 663
Qy 738 TCAGGCATATTCAAGACCTTCTCTCATGCTCGAAACTTCCAGAGCAACAATCTCTAAG 797
Db 664 TCAGTCATAATCAAAACACTGCTTCTCATGCTGGAGGCTTCCAGAGACACAGATCTCTAAG 723
Qy 798 ATCATCTCTCTTGTAGTGGTCTCTCTGCTGACCCAGACACACCTTCAACCTTGCCTATG 857
Db 724 ATCATCTCTCTTGTAGTGGTCTCTCTGCTGACCCAGATGCTTCAACCTTCAACCTTCAAG 783
Qy 858 TTAATCCAAAGTACAAGCTGGGAGTACTATACCAATCAACAGCTTTAAGTATGCCATCGTA 917
Db 784 TTCATCCGACACACACTGGGATACTATGCCATGACCACTTCTCAACCATCATG 843
Qy 918 GTGACAGAGCTATAGCATCTTTCGGCTTGGCTTAAACCTGTACTTTATGCTTTGTT 977
Db 844 GTGACAGAGGCTATGCTATCTGAGGCTTGGCTTAAACCTGTACTTTATGCTTTGTT 903
Qy 978 GCTTTAAAGTTCGGAAGACCTGCTGAACTTATGAAGATATCGGCTGCTCTCTCAC 1037
Db 904 AGCTGAAGTTCGGAAGACCTGCTGAACTTATGAAGATATCGGCTGCTCTCTCAC 963
Qy 1038 CTGGAGTCTCAAGTCAATGAAGTCTTCTGAGGACAGTTCGAAGATTTGCTGCTCC 1097
Db 964 CTTGGGCTCATACATCAATGAAGTCTTCTGAGGACATTCGAAGATTTGCTGCTCC 1023
Qy 1098 CACATGTAGAGACACCAAGTATGTTCCAAATTTGATAGGCTTGGCCACATTCAGAGAAG 1157
Db 1024 CACAAATGGAGGCCACCAAGTATGTTCCAAATTTGATAGGCTTGGCCACATTCAGAGAAG 1080
Qy 1158 TTAATAACAGAAATCTAGGAGCATGCTGTATCAATTTGGATGCAACACAGAAAAGCTTTGC 1217
Db 1081 GCTGCTCTGGAATTTGCAAGTCAATGCTGCTGCTCTTGTATGTTGAGGAGGCTTTGT 1140
Qy 1218 TTATAGCATGTGGAGTATCATGGAGAAG 1245
Db 1141 TTATAGCTTGGCATTCTCATGGAGAAG 1168

RESULT 7
AA07291
ID AAX07291 standard; cDNA; 1029 BP.
XX
AC AAX07291;
XX
DT 21-MAY-1999 (first entry)
XX
DE Pig-tailed macaque HIV/SIV receptor protein Bonzo cDNA.
XX
KW SIV receptor; HIV receptor; Bonzo; translocation promoting agent;
KW G protein coupled receptor; chemokine receptor; AIDS; infection;
KW virus transmission; pig-tailed macaque; ds.
XX
OS Macaca nemestrina.
XX
PN WO9903888-A1.
XX
```

```
PD 28-JAN-1999.
XX
PF 17-JUL-1998; 98WO-US014857.
XX
PR 17-JUL-1997; 97US-00896155.
XX
PA (UUNY ) UNIV NEW YORK STATE.
XX
PI Littman DR, Deng H, Unutmaz D, Kewalramani VN;
XX
DR WPI; 1999-132164/11.
XX
DR P-FSDE; AAW97785.
XX
PT New nucleic acid encoding the human translocation promoter Bonzo - used
PT to screen for potential agents for treating acquired immune deficiency
PT syndrome.
XX
PS Disclosure; Page 81-82; 97pp; English.
XX
CC This cDNA clone codes for pig-tailed macaque Bonzo (see AAW97785), a
CC novel HIV/SIV translocation promoting agent that acts in conjunction with
CC CD4 to serve as a receptor for the entry into a cell of a virus having a
CC specific viral envelope glycoprotein. Bonzo is a member of the 7-
CC transmembrane G-protein coupled receptor family, is closely related to
CC the chemokine receptor family, and is expressed in lymphoid tissues. The
CC invention provides the amino acid sequences (see AAW97783-88) and DNA
CC sequences (see AAX07289-94) of human, African green monkey and pig-tailed
CC macaque Bonzo and Bob (brother of Bonzo) translocation promoting agents.
CC These novel receptors were identified using an expression cloning
CC strategy. They were found to be used by particular strains of HIV-2 and M-
CC tropic HIV-1. The invention also includes: mammalian cells transfected
CC with Bonzo and/or Bob and human CD4, which can be used to screen
CC potential therapeutic agents and identify ligands; antibodies to Bonzo,
CC which can be used therapeutically, e.g. as antagonists or to target
CC toxins or radioisotopes to HIV-permissive cells; transgenic animals; and
CC antisense and ribozyme molecules, which may also be used therapeutically,
CC particularly expressed from a gene therapy vector
XX
SQ Sequence 1029 BP; 230 A; 278 C; 236 G; 285 T; 0 U; 0 Other;

Query Match 55.1%; Score 689.4; DB 2; Length 1029;
Best Local Similarity 79.8%; Pred. No. 4.2e-204;
Matches 811; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

Qy 118 ACTACGAGGAGATTTCTGCTCTTCAACAATTCACAGTATACAGCCAGGAGAAACAAAC 177
Db 14 ATTACCATGAAGACTATGGCTCAACAGTTTCATCAGCAGCCAGGAGGAGCATCAAG 73
Qy 178 GCTTCTCTAAAGTTCAAGGAGGCTTTTGGCTGTGTGTACCTGGTAGTGTGTGTCTTTG 237
Db 74 ACTTCTCTGCAAGTTCAGGAGGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 133
Qy 238 GACTGCTAGAAACTCCCTGGTTCTGATTATATACATTTTCTACAGAAAGCTGAGGACTC 297
Db 134 GCCTGTGGGGAACCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 193
Qy 298 TGACAGATGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
Db 194 TGACGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 253
Qy 358 CTTTGTGGGCTATGAGGACCTATGAGTGGGCTTCTTGGCACAGTCAATGTGCAAACTC 417
Db 254 CTTTGTGGGCTATGAGGACCTATGAGTGGGCTTCTTGGCCAGGCTATGTGCAAGACCC 313
Qy 418 TTGAGGAGCATGTATACAAATGAATCTTACGTCCTGCTGCTGCTGCTGCTGCTGCTGCT 477
Db 314 TACTGGGCTGTACACTATTAATTTCTACACATCCATGCTCATCTCCACCTGATCACTG 373
Qy 478 TGGATCCTTTTCAATGTAGTGTGCTGAGGCTACCAAGGCTTCAACCGGACGCTTAAGTGA 537
Db 374 TGGATCCTTTTCAATGTAGTGTGCTGAGGCTTCAACCGGACGCTTAAGTGAAGGA 433
Qy 538 AGATCTGGGCGCAAGTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
```

Db 434 TGACTTGGGGAGGTCTATCTGCTGCTCATCTGGGGATATCCCTGCTGTTCTTCCTG 493
 QY 598 CACAGATCATCTATGGGCATGTTCAAGATATGACAGCTTATCTGTCAGTACCAAGTG 657
 Db 494 CCCAAATATCTATGGCAATGCTTTTAACTGGACAAAGCTCATATGTTGTTATCATGACA 553
 QY 658 AGGAGATATCCACTATGTTCTGTTATACAGAGACTCTGGGGTTCTTCTGCTGCAATGC 717
 Db 554 AGGAGATTTCCACTGTTGTTCTGCCACCCAGATGACACTGGGGTTCTTCTGCTGCAATGC 613
 QY 718 TCACATGATGTTCTGTTACTCAGGCATATTAACAGACTCTGCTTCATGCTCGAAACTTCC 777
 Db 614 TCGCATGATGTTCTGCTATTCAGTCAATATCAAAACACTGCTTCATGCTGGAGGCTTCC 673
 QY 778 AGAGCACAAATCTTAAAGATCATCTTCTGTTAGTGGCTGTTGTTCTGCTGCAACGAGA 837
 Db 674 AGAAGCACAGATCTTAAAGATCATCTTCTGTTAGTGGCTGTTGTTCTGCTGCAACGAGA 733
 QY 838 CACCTTTCAACCTTGCATGTTAATCCAAAGTACAAGCTGGAGTACTATACCATTAACCA 897
 Db 734 CACCTTTCAACCTTGCATGTTAATCCAAAGTACAAGCTGGAGTACTATACCATTAACCA 793
 QY 898 GCTTTAGTATGCAATGTTAGTACAGAGGCTATAGCATCTTTCGGGCTTGCTTAAACC 957
 Db 794 GCTTTACTACCATCATATAGTACAGAGGCTATGCAATCTTTCGGGCTTGCTTAAACC 853
 QY 958 CTGTACTTTATGCTTGTGTTGCTTAAAGTTCCGGAAGACGCTCGAAACTTATCAAGG 1017
 Db 854 CTGTCTATGCTTGTGTTGCTTAAAGTTCCGGAAGACGCTCGAAACTTATCAAGG 913
 QY 1018 ATATGCTGCTCTCTACCTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTT 1077
 Db 914 ACATGCTGCTCTCTACCTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTT 973
 QY 1078 CCAGAGCTTCTGCTGCCCAATGATAGAGACCACTGATGTTCCAAATGTAG 1133
 Db 974 CCAGAGCTTCTGCTGCCCAATGATAGAGACCACTGATGTTCCAAATGTAG 1029

RESULT 8

AAS07616

ID AAS07616 standard; cDNA; 1029 BP.

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

XX AAS07616;

DR WPI; 2001-424455/45.
 DR P-PSDB; AAU04034.
 XX Identifying an agent for treating AIDS comprises administering a
 PT potential therapeutic agent to a cell expressing human CD4 and primate
 PT Bonzo or BOB and contacting the cell with a virus pseudotyped with an HIV
 PT envelope glycoprotein.
 XX Example 1; Col 43-44; 41pp; English.
 PS The sequence encodes a novel SIV (Simian immunodeficiency virus)
 XX translocating agent, Bonzo. The invention relates to selecting an agent
 CC that may be used in treating acquired immunodeficiency syndrome (AIDS)
 CC comprises administering a potential therapeutic agent to a cell
 CC expressing human CD4 and primate Bonzo or BOB (Brother of Bonzo),
 CC contacting the cell with a virus pseudotyped with an HIV envelope
 CC glycoprotein, and measuring the cell's ability to resist infection. The
 CC method is useful for selecting or identifying an agent, which can be used
 CC in treating AIDS

XX Sequence 1029 BP; 230 A; 278 C; 236 G; 285 T; 0 U; 0 Other;
 SQ

Query Match 55.1%; Score 689.4; DB 5; Length 1029;
 Best Local Similarity 79.8%; Pred. No. 4.2e-204;

Matches 811; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

QY 118 ACTACGAGGGAGATTCTCGGCTCTTCAACATTCAGTATACAGCCAGGAGAACAAAC 177
 Db 14 ATTACATGAAGACTATGGGCTCAACAGTTTCAATGACAGCCAGGAGAGCATCAAG 73
 QY 178 GCTTCCTAAAGTTCAGAGAGGTCTTTTGGCCCTGTGTACCTGGTAGTGTGTCTTTG 237
 Db 74 ACTTCCTGCACTTCAGGAAGGTCTTCTGCTGCTGATGTACCTGGTGTGTGTCTGTG 133
 QY 238 GACTGCTAGGAACCTCCCTGCTGCTGATATATACATTTTACCAGAACTGAGGACTC 297
 Db 134 GCCTGGTGGGAACTCCCTGCTGCTGATATCCATCTTCTACCATAACTGCGAGGCC 193
 QY 298 TGACAGATGTCTTCTGCTGAACTTGGCCCTGGCTGACCTGGTGTGTGTGTGTGTGTG 357
 Db 194 TGACGGAGCTGTCTGCTGTAACCTACCCCTGGCTGACCTGGTGTGTGTGTGTGTGTG 253
 QY 358 CTTTGGGCTTATGAGGACCTATGAGTGGGTCTTTGGGACAGTATGTGCAAACTC 417
 Db 254 CTTTGGGCTTATGAGGACCTATGAGTGGGTCTTTGGGACAGTATGTGCAAACTC 313
 QY 418 TTGAGGAGCATGTATACAAATGAATCTTACGTGTCTGCTGCTGCTGCTGCTGCTGCTG 477
 Db 314 TACTGGGCTTACACTATTAATCTTACATCCATGCTGCTGCTGCTGCTGCTGCTGCTG 373
 QY 478 TGGATCGTTTCATTTAGTGTCCAGGCTACAGGCTTCAACGGCTTCAACGGCAGGTAAGTGA 537
 Db 374 TGGATCGTTTCATTTAGTGTCCAGGCTTCAACGGCTTCAACGGCAGGTAAGTGA 433
 QY 538 AGATCTGGGCAAGTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
 Db 434 TGACTTGGGCAAGTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 493
 QY 598 CACAGATCATCTATGGCCATGTTCAAGATATGACAAAGCTTATCTGTCACTGACACAGTG 657
 Db 494 CCCAAATATCTATGGAATGTCTTTAATCTGGACAAGCTCATATGTGTTATCATGACA 553
 QY 658 AGGAGATATCCACTATGGTCTTGTATACAGATGACTCTGGGGTCTTCTGCTGCTGCTG 717
 Db 554 AGGAGATTTCCACTGTTGTTCTTGGCACCCAGATGACACTGGGGTCTTCTTCCCACTGC 613
 QY 718 TCACATGATTTCTGCTACTCAGGCAATATCAAGACTTTCATGCTGCTGCTGCTGCTGCT 777
 Db 614 TCGCATGATTTCTGCTATTCAGTCAATCAAAACACTGCTTTCATGCTGAGGCTTCC 673
 QY 778 AGAAGCAAAATCTTAAAGATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
 Db 674 AGAAGCAAGATCTTAAAGATCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 733

Littman DR, Deng H, Unutmaz D, Kewalramani VN;

QY 838 CACCTTCACCTTGCATGTTAATCCAAAGTACACAGCTGGAGTACTATACCAACCA 897
DB 734 CACCTTCACCTTGCATGTTAATCCAAAGTACACAGCTGGAGTACTATACCAACCA 793
QY 898 GCTTTAAGTATGCTATGCTAGTACAGAGGCTATAGCATATCTTCCGGCTTGGCTTAACC 957
DB 794 GCTTTAAGTATGCTATGCTAGTACAGAGGCTATAGCATATCTTCCGGCTTGGCTTAACC 853
QY 958 CTCTACTTATGCTTGTGCTTAAAGTTCGGAGAGACCTCTGGAACCTTATGAAGG 1017
DB 854 CTCTACTTATGCTTGTGCTTAAAGTTCGGAGAGACCTCTGGAACCTTATGAAGG 913
QY 1018 ATATCGGCTGCTCTCTACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACATTT 1077
DB 914 ACATGCTGCTCTCTCTACCTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACATTT 973
QY 1078 CCAAGCTTGTCTGCTCTCCCAATGATGAGACACCAAGTATGTTCAATGTTAG 1133
DB 974 CCAAGCTTGTCTGCTCTCCCAATGATGAGACACCAAGTATGTTCAATGTTAG 1029

RESULT 9
ID AEN84641
AC AEN84641 standard; cDNA; 1029 BP.
XX AEN84641;
XX 28-OCT-2002 (first entry)
XX Pig-tailed macaque G-protein coupled receptor Bonzo cDNA.

XX Bonzo; G-protein coupled receptor; receptor; STRL33; TYMSTR; retrovirus;
XX translation promoting agent; anti-HIV; antiviral; virucide; AIDS;
XX antiinflammatory; immunostimulant; transgenic animal; pig-tailed macaque;
XX gene; ss.
XX Macaca remestrina.

XX Key Location/Qualifiers
XX CDS 1..1029
XX /*tag= a
XX /product= "Bonzo"
XX US2002076694-A1.
XX 20-JUN-2002.
XX 09-MAY-2001; 2001US-00852156.
XX 17-JUL-1997; 97US-0052827P.
XX 16-JUL-1998; 98US-00116498.

XX (LITT/) LITTMAN D R.
XX (DENG/) DENG H.
XX (UNUT/) UNUTMAZ D.
XX (KEMA/) KEWALRAMANI V N.
XX Littman DR, Deng H, Unutmaz D, Kewalramani VN;
XX WPI; 2002-598719/64.
XX P-PSDB; ABB79745.

XX New isolated nucleic acid encoding the translocation promoting agent
XX Bonzo, useful e.g. in screening for agents that inhibit infection by
XX human immune deficiency virus.
XX Disclosure; Page 31; 65pp; English.

XX The present sequence is the coding sequence for pig-tailed macaque Bonzo
XX (also named STRL33 or TYMSTR). Human Bonzo (see ABB79743) has been
XX identified as a novel G-protein coupled receptor involved in retroviral
XX entry into cells, acting in conjunction with CD4 as a translocation

CC promoting agent for SIV and also for certain strains of HIV-2 and M-
CC tropic HIV-1. Bonzo and the closely related BOB (brother of Bonzo) are
CC related to the chemokine receptor family and are expressed in lymphoid
CC tissues. Use of these new receptors following experimental infection of
CC non-human primates with SIV strains may provide important insight into
CC viral transmission and mechanisms of SIV- and HIV-induced AIDS. The
CC invention provides Bonzo and BOB amino acid and DNA sequences, mammalian
CC cells transfected with Bonzo and/or BOB and human CD4, and antibodies to
CC Bonzo. A method of identifying other translocation promoting agents is
CC also disclosed. Cells (or transgenic animals) that express Bonzo or BOB
CC and CD4 are used to: (a) identify therapeutic agents for treating AIDS
CC (acquired immune deficiency syndrome), by increasing the resistance of
CC cells to infection; (b) in filters to remove viruses with HIV envelope
CC glycoproteins from biological fluids; (c) to identify ligands for Bonzo
CC or BOB; (d) to identify compounds that increase the immune response to a
CC specific pathogen or vaccine by increasing Bonzo/STRL33 expression; and
CC (e) to identify antiinflammatory agents that reduce Bonzo/STRL33
CC expression and thus inhibit recruitment of memory cells
XX
XX SQ Sequence 1029 BP; 230 A; 278 C; 236 G; 285 T; 0 U; 0 Other;

Query Match 55.1%; Score 689.4; DB 6; Length 1029;
Best Local Similarity 79.8%; Pred. No. 4.2e-204;
Matches 8.1; Conservative 0; Mismatches 205; Indels 0; Gaps 0;
QY 118 ACTACGAGGAGATTCTGGCTCTTCAACAATTCAGTGTATAACAGCAGAGAACAAAC 177
DB 14 ATTACCATGAAGACTATGGGCTCAACAGTTTCAATGACAGCAGCAGGAGGAGCATCAAG 73
QY 178 GCTTCTCTAAAGTTCAAGGAGGTTCTTTTGGCTGTGTACTGTGTACTGTGTGTCTTTG 237
DB 74 ACTTCTCTGAGTTCAGGAAGGTTCTTTTGGCTGTGTACTGTGTGTGTGTGTGTGTG 133
QY 238 GACTGTCTAGGAACCTCCCTGGTCTGTGATATATATATATTTTACAGAGCTGAGGACTC 297
DB 134 GCCTGTGTGGGAACTCCCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 193
QY 298 TGACAGATG 357
DB 194 TGACGAGAGTG 253
QY 358 CTTTGTGGGCTTATGAGGACCTATGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 417
DB 254 CTTTGTGGGCTTATGAGGACCTATGAGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 313
QY 418 TTGAGGAGATGTATACAAATGAATCTTACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 477
DB 314 TACTGGGCTGTACACTATTAATCTTACATCAATCAATCAATCAATCAATCAATCAATCA 373
QY 478 TGGATGTTTCAATGTAGTG 537
DB 374 TGGATGTTTCAATGTAGTG 433
QY 538 AGATCTGGGGCAAGTCAATTTG 597
DB 434 TGACTTGGGGCAAGTCAATTTG 493
QY 598 CACAGATCATCTATGGCCATGTTCAAGATATTGACAAGCTTATCTGTACGTACACAGTG 657
DB 494 CCAAAATTTATCTATGGCAATGTTTAACTTGAACAAGCTTATCTGTACGTACACAGTG 553
QY 658 AGGAGATATCCACTATGTTCTG 717
DB 554 AGGAGATATCCACTATGTTCTG 613
QY 718 TCATATGATTTG 777
DB 614 TGGCCATGATTTG 673
QY 778 AGAAGCACAATCTCTAAAGATCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 837
DB 674 AGAAGCACAATCTCTAAAGATCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 733

Qy	1097	CCACAATGTAGAGACCACCAGTATGTTCCAATTGTAG	1133
Db	993	CCACAATGTGGAGGCCACGAGATGTTCCAGTTATAG	1029

RESULT 11

AAS07615
 ID AAS07615 standard: cDNA: 1037 BP.

AC AAS07615:

23-OCT-2001 (first entry)

XX DE Green monkey cdNA encoding the SIV translocation agent, Bonzo.

AA Monkey; Bonzo; simian immunodeficiency virus; SIV translocating agent;
KW CD4; B05; Brother of Bonzo; acquired immunodeficiency syndrome; AIDS;
KW SIV; HIV; ss.

Cercopithecus aethiops.

	Key	Location/Qualifiers
FT	CDS	1. .1029
FT		/*tag= a
FT		/product= "Bonzo"

YY
PN
1156251582-B1

XX
DN
26-JUN-2001

XX
PF
,
16-JUL-1998.
98UIS-00116498

XX
PB 17-III-1997: 97IIS-0052827P

XX
P2 (UNY) INIV NEW YORK STATE

XX
PI Rittman DP Peng H Umutaz D Kewaj ramani VN:

XX
DR WPI; 2001-424455/45.
DR P-PSDE: AAU04033.

Identifying an agent for treating AIDS comprises administering a potential therapeutic agent to a cell expressing human CD4 and primate Benvo or BOB and contacting the cell with a virus pseudotyped with an HIV envelope glycoprotein.

XX
PS Example 1: Col 39-40; 41pp; English.

The sequence encodes a novel HIV (Simian immunodeficiency virus) translocating agent, Bonzo. The invention relates to selecting an agent that may be used in treating acquired immunodeficiency syndrome (AIDS) comprising administering a potential therapeutic agent to a cell expressing human CD4 and primate Bonzo or BOB (Brother of Bonzo), contacting the cell with a virus pseudotyped with an HIV envelope, glycoprotein, and measuring the cell's ability to resist infection. The method is useful for selecting or identifying an agent, which can be used in treating AIDS.

Sequence 1037 BP: 232 A; 276 C; 235 G; 294 T; 0 U; 0 Other; XX SO

Query Match 54.8%; Score 685; DB 5; Length 1037;
Best Local Similarity 81.1%;
Pred. NO. 9.9e-203;
Matches 809; Conservative 0; Mismatches 185; Indels 3; Gaps 17

140 CTTCAACAATTCCAGTGATAACAGCCAGGAGAACAAAC--GCTTCCCTAAAGTTCAAGGA 196

D_b 33 CTTCAACAGTTTCAATGACAGCAGCCAGGAGCATCAGGACTTCCTGCAGTTCAGCAA 92

XX African green monkey G-protein coupled receptor Bonzo cDNA.
DE
XX
XX Bonzo; G-protein coupled receptor; receptor; STRL33; TMSTR; retrovirus;
KW translation promoting agent; anti-HIV; antiviral; virucide; AIDS;
XX

KW myocardial infarction; MI; stroke; glaucoma; anxiety;
 KW prostatic hyperplasia; epilepsy; prostate cancer; rhinitis; angina;
 KW prostatic hypertrophy; gene; ss.
 XX Homo sapiens.
 XX WO300268600-A2.
 XX PD 06-SEP-2002.
 XX PF 26-FEB-2002; 2002WO-US005625.
 XX PR 26-FEB-2001; 2001US-0271913P.
 XX PA (AREN-) ARENA PHARM INC.
 XX PI Liaw CW, Chalmers DT, Behan DP, Maciejewski-Lenior D, Leonard JN;
 PI Lin I, Ortuno D;
 XX WPI; 2002-706980/76.
 DR P-PSDB; ABG95153.
 XX New human G-protein coupled receptor (GPCR), useful for screening agonist
 PT or inverse agonist compounds for treating diseases associated with GPCR.
 XX Claim 7; Page 107-108; 201pp; English.
 XX The present invention relates to transmembrane receptors, particularly
 CC endogenous human G-protein coupled receptors (GPCRs), mutant (non-
 CC endogenous) versions of the GPCRs, and the polynucleotide sequences
 CC encoding them. The GPCRs are useful for screening agonist or inverse
 CC agonist compounds for treating diseases associated with GPCR. Diseases
 CC that can be treated with such compounds include allergies, hypertension,
 CC reflux disease, depression, migraine, schizophrenia, ulcers, psychotic
 CC disorders, asthma, bronchospasm, anaesthesia, myocardial infarction (MI),
 CC stroke, glaucoma, prostatic hyperplasia, epilepsy, prostate cancer,
 CC anxiety, prostatic hypertrophy, rhinitis, and angina. The present
 CC sequence encodes an endogenous human GPCR
 XX Sequence 1029 BP; 229 A; 276 C; 237 G; 287 T; 0 U; 0 Other;

Query Match 54.4%; Score 680.8; DB 6; Length 1029;
 Best Local Similarity 81.3%; Pred. No. 2e-201;
 Matches 787; Conservative 0; Mismatches 181; Indels 0; Gaps 0;
 QY 166 AGGAGAACAAACGCTTCTTAAGTTCAAGGAGGCTTTTGGCCCTGTGTGACCTGGTAG 225
 DB 62 AGGAGCATCAAGCCTTCTGTCAGTTTCAGCAAGGCTTTCTGCTGCTGACCTGGTGG 121
 QY 226 TGTGTTGCTTTGGACTGTAGGAACTCCCTGGTCTGATTATATACATTTTCTACCCAGA 285
 DB 122 TGTGTTGCTGTTGGTGTGGGAACTCTCTGGTCTGTTATCCATCTTCTACCATTA 181
 QY 286 AGCTGAGGACTGTGACAGATGTTTCTGCTGAATTCGCCCTGGCTGACCTGGTGTGTTG 345
 DB 182 AGTTGACAGGCTGACGAGATGTTTCTGCTGTAACCTACCCCTGGCTGACCTGGTGTGTTG 241
 QY 346 TCTGTACTCTGCCCTTTTGGGCTATGAGGACCTATGAGTGGTCTTTGSCACAGTCA 405
 DB 242 TCTGCACITCGCCCTTCTGGGCTATGAGGACCTATGAGTGGTGTGTTGGCCAGGTCA 301
 QY 406 TGTGCAAAACTCTTCGAGGCATGTATACAAATGAATTTCTAGCTGTCCATGTCTCACTCA 465
 DB 302 TGTGCAAAAGCCTACTGGSCATCTACACTATTAACTTCTACAGCTCCATGTCTCACTCA 361
 QY 466 CTGTGATCACTGATGCTTTTCATGTAGTGGTCCAGGCTACCAAGGCTTCAACGGC 525
 DB 362 CTGTGATCACTGATGCTTTTCATGTAGTGGTAAAGGCCACCAAGGCTTCAACAGC 421
 QY 526 AGGCTAAGTGGAGATCTGGGGCCCAAGTCATTTGCTGCTATTTGGTGGTGTCTCCCTGT 585
 DB 422 AAGCCAAAGGATGACCTGGGGCAAGGTCACCGATTTGCTCATCTGCTGGTGTATCCCTGC 481

QY 586 TGTGTTCTTTGGCAGAGATCACTATATGGCAATGTTCAAGATATATGACAAGCTTATCTGTC 645
 DB 482 TGTGTTCTTTGGCAGAGATCACTATATGGCAATGTTCTTTAATCTCGACAAGCTCATATGTG 541
 QY 646 AGTACCACAGTGGAGAGATATCACTATATGTTCTTTGTTATACAGATGACTCTGGGGTTCT 705
 DB 542 GTTACCATGACGAGGCAATTTCCACTGTGTGTTCTTCCACCCAGATGACACTGGGGTTCT 601
 QY 706 TCTGTCATTTGCTCACTATGATTTCTGTCTACTCAGGCATTTATCAAGACTTGTGTTCTATG 765
 DB 602 TCTTGCACCTGCTCACCATGATTTGTCTGCTATTCACTCATATCAAAACACTGCTTCATG 661
 QY 766 CTGCAAACTTCCAGAGCACAATCTCTAAAGATCATCTTCTTGTAGTGGTGTGTTCC 825
 DB 662 CTGGAGGCTTCCAGAGCAGATCTCTAAAGATCATCTTCTGTTGGTGTGTTGTTCC 721
 QY 826 TGTGACCCAGACACCCCTTCAACCTTGGCATTTGATCCAAAGTACAAGCTGGGAGTACT 885
 DB 722 TGTGACCCAGATGCTTCAACCTCATGAGTTTCATCCGACGACACACTGGGATACT 781
 QY 886 ATACCATAACAGCTTTAAGTATGTCATCGTATGACAGAGGCTATAGCATATTTTCGGG 945
 DB 782 ATGCCATGACAGCTTTTCACTACACCATCATGCTGACAGAGGCCATCGCATACCTGAGG 841
 QY 946 CTGTCCTTAACCTGACTTTTATGCTTGTGGCTTAAAGTTCCGGAAGACGCTCTGA 1005
 DB 842 CTGTCCTTAACCTGCTGCTATGCTTGTGACGCTGAGTTTCGAAGAACTTCTGGA 901
 QY 1006 AACTTATGAAGATATCGGCTGCTCTCTCACTGGGAGTCTCAAGTCAATGAAATCTT 1065
 DB 902 AACTTGTGAAGACATTTGGTTGGCTCCCTTACCTGGGGTCTCACATCAATGAAATCTT 961
 QY 1066 CTGAGCAGCTTCCAGAGCTTGTCTGCTCCCAATGTAGAGACCCAGCATCTCTCC 1125
 DB 962 CTGAGCAGCTTCCAGAGCTTGTCTGCTCCCAATGTAGAGACCCAGCATCTCTCC 1021
 QY 1126 AATTGTAG 1133
 DB 1022 AGTTATAG 1029
 RESULT 14
 AAF85364
 ID AAF85364 standard; cDNA; 1029 BP.
 XX AAF85364;
 AC AAF85364;
 XX 23-JUL-2001 (first entry)
 DT 23-JUL-2001 (first entry)
 XX Nucleotide sequence of a human Bonzo polypeptide.
 DE Bonzo; CXK chemokine receptor; inflammatory disease; cancer; infection;
 KW SEXKline; spleen extracted chemokine; ss.
 XX Homo sapiens.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 1..1029
 FT /tag= a
 FT /product= "Bonzo"
 XX WO200137872-Al.
 XX 31-MAY-2001.
 XX 22-NOV-2000; 2000WO-US032206.
 XX 24-NOV-1999; 99US-00449437.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Briskin MJ, Murphy KE, Wilbanks AM, Wu L;
 XX

DR WPI; 2001-343947/36.
 DR P-PSDB; AAB69422.
 XX
 PT Identifying agents (especially antibodies) which bind to the CXCR
 PT chemokine receptor Bonzo, and which may be used to treat e.g. cancers and
 PT inflammation.
 XX
 PS Example 1; Fig 1; 191pp; English.
 XX
 CC The present sequence encodes a human Bonzo polypeptide. Bonzo is a CXCR
 CC chemokine receptor. The specification describes a method for identifying
 CC agents (especially antibodies) which bind to Bonzo and inhibit the
 CC binding of a ligand (especially CXCR1) (splenectin extracted chemokine))
 CC and the agents per se. The agents identified may be used for the
 CC treatment of a disorder/disease related to aberrant Bonzo expression and
 CC activity, such as inflammatory disease, cancers and/or infections (e.g.
 CC viral, bacterial and fungal infections)
 XX
 SQ Sequence 1029 BP; 229 A; 275 C; 238 G; 287 T; 0 U; 0 Other;

Query Match 54.3%; Score 679.2; DB 4; Length 1029;
 Best Local Similarity 81.2%; Pred. No. 6.4e-201;
 Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 166 AGGAGAACAAACGGCTTCTTAAGTTTCAAGGAGGCTTTTGGCCCTGTGTGCTACCTGGTAG 225
 DB 62 AGGAGCATCAAGACTTCTCCAGTTCAGCAAGGCTTTTCTGCCCTGCATGTACCTGGTGG 121
 QY 226 TGTTCCTTTGGAGCTAGGAACCTCCCTGGTTCGTGATATATATATATATATATATATAT 285
 DB 122 TGTTCCTTTGGAGCTAGGAACCTCCCTGGTTCGTGATATATATATATATATATATATAT 181
 QY 286 AGCTGAGGACTCTGACAGATGTCTCTGCTGAATTTGCCCTGGCTGACCTGGTGTGTTG 345
 DB 182 AGTTGAGAGGCTGACGAGATGTCTCTGCTGAATTTGCCCTGGCTGACCTGGTGTGTTG 241
 QY 346 TCTGTACTTGGCCCTTTGGGCTATGACGACCTATGAGTGGTCTTTGGGACACAGTCA 405
 DB 242 TCTGCACTTGGCCCTTTGGGCTATGACGACCTATGAGTGGTCTTTGGGACACAGTCA 301
 QY 406 TGTCAAAACCTTTCGAGGAGATGATACATGATGATGATGATGATGATGATGATGATGATG 465
 DB 302 TGTCAAAACCTTTCGAGGAGATGATACATGATGATGATGATGATGATGATGATGATGATG 361
 QY 466 CTTGCATCAGAGTGGATCGTTTCATGAGTGGTCCAGGCTACCAAGGCTTCAACCGGC 525
 DB 362 CTTGCATCAGTGGATCGTTTCATGAGTGGTCCAGGCTACCAAGGCTTCAACCGGC 421
 QY 526 AGGCTAAGTGGAGATCTGGGCGCAAGTCAATTTGCTGCTCATTTGGTGGTCTCCCTGT 585
 DB 422 AAGCCAAAGAGATGACCTGGGCGCAAGTCAATTTGCTGCTCATTTGGTGGTCTCCCTGT 481
 QY 586 TGGTTTCTTTGCCACAGATCATCTATGGCCATGTTTCAAGATATTGCAAGCTTATCTGTC 645
 DB 482 TGGTTTCTTTGCCACAAATTTATCTATGGCAATGTTTCAATTCGCAAGCTTATCTGTC 541
 QY 646 AGTACACAGTGGAGATATCCACTATGTTTCTGTTATACAGATGACCTGGGTTCT 705
 DB 542 GTTACCATGAGGAGCAATTTCCACTGTGTTCTTGGCCACCCAGATGACACTGGGTTCT 601
 QY 706 TCTGTCATGCTCAGATCATCTGTTGCTACTCAGGCAATTTCAAGCTTATCTGTCATG 765
 DB 602 TCTGTCATGCTCAGATCATCTGTTGCTACTCAGGCAATTTCAAGCTTATCTGTCATG 661
 QY 766 CTCGAATCTTCCAGAGACAAATCTCTAAGATCATCTTCTTGTAGTGGTGTGTTCTTC 825
 DB 662 CTGAGGCTTCCAGAGACAGATCTCTAAGATCATCTTCTTGTAGTGGTGTGTTCTTC 721
 QY 826 TGTGACCCAGACACCTTCAACCTTCCCATGTTTATCCAAAGTACCAAGCTGGAGTACT 885
 DB 722 TGTGACCCAGATGCTTCAACCTTCCCATGTTTATCCAAAGTACCAAGCTGGAGTACT 781
 QY 886 ATACCATACCAAGCTTTTAAATGATGCCATCGTAGGACAGGCTATAGCATACTTTCCGG 945

DB 782 ATGCCATACACAGCTTTTCTACATACATCATGTGACAGAGCCATCCATACCTGAGGG 841
 QY 946 CTTGCTTTAAACCTTACTTATGCTTTTGTGGCTTAAAGTTCCGGAAGAACCTGTGA 1005
 DB 842 CTTGCTTTAAACCTTACTTATGCTTTTGTGGCTTAAAGTTCCGGAAGAACCTGTGA 901
 QY 1006 AACTTATGAAGATATCGGCTGCTCTCTACCTGGGAGTCTCAAGTCAATGGAAGTCTT 1065
 DB 902 AACTTGTGAAGGACATGTTGGCTTCCCTTACCTTGGGGTCTCATCATCAATGGAATCTT 961
 QY 1066 CTGAGGACAGTTTCCAGACTTGTCTGCTCCCAATGTAGAGACCAACCATGATGTTCC 1125
 DB 962 CTGAGGACAAATTCAGACTTTTCTGCTCCCAATGTAGAGGACCAACCATGATGTTCC 1021
 QY 1126 AATTGTAG 1133
 DB 1022 AGTTATAG 1029

RESULT 15
 ADE31648
 ID ADE31648 standard; DNA; 1029 BP.
 XX
 AC ADE31648;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human 6193 gene #SEQ ID 5.
 XX
 KW Antiarteriosclerotic; cardiatic; vasotropic; antiinflammatory;
 KW thrombolytic; antiarrhythmic; antianaginal; hypotensive; gene therapy;
 KW cardiovascular; disorder; ischaemia; aortic bending;
 KW vascular heart disease; endocarditis; atrial fibrillation; heart failure;
 KW angina; cardiomyopathy; cardiac death; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003065984-A2.
 XX
 PD 14-AUG-2003.
 XX
 PF 29-JAN-2003; 2003WO-US002571.
 XX
 PR 01-FEB-2002; 2002US-0353224P.
 PR 15-MAR-2002; 2002US-0364529P.
 PR 19-APR-2002; 2002US-0373861P.
 PR 29-APR-2002; 2002US-0376287P.
 PR 12-JUN-2002; 2002US-0388080P.
 PR 24-JUN-2002; 2002US-0390971P.
 PR 03-JUL-2002; 2002US-0394130P.
 PR 10-JUL-2002; 2002US-0394797P.
 PR 21-AUG-2002; 2002US-0405450P.
 PR 04-SEP-2002; 2002US-0408070P.
 PR 06-NOV-2002; 2002US-0424300P.
 PR 05-DEC-2002; 2002US-0431042P.
 PR 05-DEC-2002; 2002US-0431079P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Logan TJ, Chun M, Galvin KM, Healy A, Acton SL, Donaghue M;
 PI Stagliano N, Perodin J, Rodrigue-Way A;
 XX
 DR WPI; 2003-731468/59.
 DR P-PSDB; ADE31649.
 XX
 PT Identifying a compound capable of treating a cardiovascular disorder
 PT (e.g. atherosclerosis) comprises assaying the ability of the compound to
 PT modulate the expression or activity of e.g. 1682, 6169 or 6193
 PT polypeptide or nucleic acid.
 XX
 PS Disclosure; SEQ ID NO 5; 328pp; English.

XX CC The invention relates to a method for identifying a compound capable of
CC treating a cardiovascular disorder. The present invention identifies the
CC differential expression of 1682, 6169, 6173, 7751, 14395, 29002, 33216,
CC 43726, 69232, 21656, 32427, 2402, 7747, 1720, 9151, 60491, 1371, 7077,
CC 33207, 1419, 18036, 16105, 38650, 14245, 58848, 1870, 25856, 32394, 3484,
CC 345, 9252, 9135, 10532, 18610, 8165, 2448, 2445, 64624, 84237, 8912,
CC 2868, 283, 2534, 9464, 17799, 26686, 43848, 32135, 12208, 2914, 51130,
CC 19489, 21833, 2947, 59590, 15992, 2094, 2252, 3474, 9792, 15400, 1452 or
CC 6585 genes in cardiovascular disease states. The methods are useful in
CC diagnosing, preventing and treating cardiovascular disorders, such as
CC atherosclerosis, cardiac hypertrophy, ischaemia, reperfusion injury,
CC restenosis, arterial inflammation, vascular wall remodeling, coronary
CC microembolism, tachycardia, bradycardia, pressure overload, aortic
CC bending, coronary artery ligation, vascular heart disease, valvular
CC disease, including but not limited to, valvular degeneration caused by
CC calcification, rheumatic heart disease, endocarditis, or complications of
CC artificial valves; atrial fibrillation, long-QT syndrome, congestive
CC heart failure, sinus node dysfunction, angina, heart failure,
CC hypertension, atrial fibrillation, atrial flutter, pericardial disease,
CC including but not limited to, pericardial effusion and pericarditis;
CC cardiomyopathies, e.g. dilated cardiomyopathy or idiopathic
CC cardiomyopathy, myocardial infarction, coronary artery disease, coronary
CC artery spasm, ischaemic disease, arrhythmia, sudden cardiac death and
CC cardiovascular developmental disorders. The methods may also be used for
CC identifying compounds that modulate cardiovascular disorders. Sequences
CC given in ADE31644-ADE31769 represent the genes and proteins that may be
CC regulated by a compound of the invention.

Sequence 1029 BP; 229 A; 275 C; 238 G; 287 T; 0 U; 0 Other; ;

Query Match 54.3%; Score 679.2; DB 9; Length 1029;
Best Local Similarity 81.2%; Pred. No. 6.4e-201;
Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

166	AGGAGAACAAACGCTCTCTAAAGTTCAAGAGGCTCTTTTGGCCCTGTGTGTACCTGGTAG	225
Db		
62	AGGAGCATCAAGACTTCTGCAGTTCGACAGGCTCTTTTGGCCCTGCATGTACCTGGTGG	121
Qy		
226	TGTTGTCTTTGGACTGCTAGGAACTCCCTGGTCTGTGATTATACATTTTCTACACAGA	285
Db		
122	TGTTTGTCTG:GGTCTGTGGGAACTCTCTGGTGTGGTCAATCCATCTTCTACCATTA	181
Qy		
286	AGCTGAGGACTCTGACAGATGTTTCTGTGAACTTGGCCCTGGCTGACCTGTGTGTTTG	345
Db		
182	AGTTGAGAGCGTGAAGGATGTGTTCTGTGTGAACCTACCCCTGGGTGACCTGGTGTTG	241
Qy		
346	TCGTACTCTGCCCTTTTGGGCTATGACAGGACCTATGATGGGTGGTCTTTTGGCACAGTCA	405
Db		
242	TCTGCACCTCTGCCCCCTCTGGGCCCTATGAGGATCCATGAATGGGTGTTTGGCCAGTGCA	301
Qy		
406	TGTGCAAAACTCTTGGAGGCATGTATACAATGAACCTTCTACGTGCCATGCTCACTCTCA	465
Db		
302	TGTGCAAGAGCCTACTGGGCCATCTACACTATTAACTTCTACACGTCCATGCTCATCCTCA	361
Qy		
466	CCTGCATCACAGTGGATCGTTTCAITGTAGTGGTCAAGGCTACCAAGGCCCTTCACCCGC	525
Db		
362	CCTGCATCACTGTGATCGTTTCAITGTAGTGGTAAAGGCCACCAAGGCCCTACACCCAGC	421
Qy		
526	AGGCTAAGTGGGAAGATCTGGGGCCCAAGTCATTTGCTTTGCTCATTTGGGTGGTCTCCCTGT	585
Db		
422	AGGCCAAGAGGATGACCTGGGGCAAGTCCACGCTTGCTCATCTGGGTGATATCCCTGC	481
Qy		
586	TGTTTCTTTGGCCA CAGATCACTATATGGCCATGTTCAAGATATATGACAAGCTTATCTGTCT	645
Db		
482	TGTTTCTCTTGGCCCCAAATATCTATGGCAAATGCTTTTAATCTCGACAAGCTCATATGTG	541
Qy		
646	AGTACCACAGTCAGGAGATATCCACTATGGTCTTTGTTATACAGATGATCTCTGGGGTCTCT	705
Db		
542	GTTACCATGACAGGCAATTTCCACTGTGGTCTTGGCCACCCAGATGACACTGGGGTCTCT	601
Qy		
706	TCTGCTAATTGCTCACTATGATTTCTGTGCTACTCAGGCATTTATCAAGACCTTGCTTCATG	765

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2004, 00:32:09 ; Search time 3116 Seconds

(without alignments)
11988.956 Million cell updates/sec

Title: US-10-603-141-1

Perfect score: 1251

Sequence: 1 gaattcgacagaggcaga.....gatatcgaggagcctgac 1251

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_nam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gsl1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1231.4	98.4	1791	11 AK052901	AK052901 Mus muscu
2	1029	82.3	1029	29 AY399296	AY399296 Mus muscu
3	859	68.7	991	13 BY746127	BY746127 BY746127
4	679.2	54.3	1029	29 AY399294	AY399294 Homo sapi

5	679.2	54.3	1029	29	AY399295	AY399295 Pan trogl
6	662	52.9	672	13	BY746487	BY746487 BY746487
7	658.2	52.6	663	14	CF912719	CF912719 A0638C02
8	647	51.7	647	14	CF911440	CF911440 A0615D03
9	631	50.4	631	14	CF913009	CF913009 A0641F09
10	627	50.1	627	14	CF911103	CF911103 A0610D08
11	626.4	50.1	628	14	CF912382	CF912382 A0631A05
12	617.4	49.4	619	14	CA574451	CA574451 K0616E05
13	616	49.2	626	10	BB662193	BB662193 BB662193
14	610	48.8	610	14	CF910757	CF910757 A0647A08
15	599.4	47.9	601	14	CF913351	CF913351 A0604D11
16	596.2	47.7	601	14	CF911739	CF911739 A0620D02
17	594.4	47.5	596	14	CA575926	CA575926 K0637E06
18	587.5	47.0	594	13	BQ562713	BQ562713 H4077E09
19	584.4	46.7	586	14	CF911576	CF911576 A0617F05
20	575.6	46.0	773	12	BG963872	BG963872 G0282855
21	572.6	45.8	579	14	CA574155	CA574155 K0611B04
22	567.4	45.4	569	14	CA576578	CA576578 K0646H04
23	558.4	44.6	560	14	CF912084	CF912084 A0626A08
24	555.8	44.4	629	14	CF910917	CF910917 A0606H11
25	549.2	43.9	554	14	CF912157	CF912157 A0627C05
26	549	43.9	608	10	BB660928	BB660928 BB660928
27	539	43.1	659	10	BB630116	BB630116 BB630116
28	514	41.1	884	12	BG963267	BG963267 G0282836
29	504.4	40.3	506	14	CA575464	CA575464 K0631A11
30	492	39.3	514	14	CF911881	CF911881 A0622G04
31	484	38.7	484	14	CA575959	CA575959 K0637H11
32	482.4	38.6	484	14	CA573379	CA573379 K0600C07
33	482.4	38.6	484	14	CA574776	CA574776 K0621A04
34	462	36.9	462	14	CF911521	CF911521 A0616F03
35	452.8	36.2	467	14	CA574254	CA574254 K0613E05
36	438.4	35.0	440	14	CA576368	CA576368 K0643C07
37	394	31.5	416	14	CA576715	CA576715 K0649B10
38	391.6	31.3	1060	13	BX398345	BX398345 BX398345
39	385.8	30.8	1201	13	BX398107	BX398107 BX398107
40	381.2	30.5	650	12	BI837704	BI837704 G03086564
41	366.6	29.3	622	9	AI648943	AI648943 uk333b08.x
42	357.4	28.6	368	13	BY214569	BY214569 BY214569
43	352	28.1	679	28	BZ263559	BZ263559 CH230-254
44	347	27.7	356	13	BY221709	BY221709 BY221709
45	345	27.6	354	13	BY221071	BY221071 BY221071

ALIGNMENTS

RESULT 1
AK052901
LOCUS
DEFINITION
Mus musculus 16 days neonate heart cDNA, RIKEN full-length enriched library, clone:D830030f09 product:chemokine (C-X-C) receptor 6, full insert sequence.
1791 bp mRNA linear HTC 20-SEP-2003
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED

AK052901
AK052901.1 GI:26343078
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Mech. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipette sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20330913
11076861

TITLE
JOURNAL
MEDLINE
PUBMEDREFERENCE
AUTHORS

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 403, 685-690 (2001)

TITLE
JOURNAL
REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1791)

TITLE
JOURNAL
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, K., Saichon, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL

Direct Submission
Submitted (15-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES
source

Location/Qualifiers
1. 1791
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:D830030F09"
/db_xref="MGI:2423199"
/db_xref="taxon:10090"
/clone="D830030F09"
/tissue_type="heart"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
73. 1128
/note="unnamed protein product; chemokine (C-X-C) receptor
6 (MGD|MGF:1934582, GB|NM_030712, evidence: BLASTN, 99%, match=1788)
/codon_start=1
/protein_id="BAG35196.1"
/db_xref="GI:26343079"

CDS

/translation="MDDGHQBSALYDGHYEGDFWLFNNSDNSENKRFLEKFEVFLP
CWLTVFVFGLLGNLVLIIYFYOKLTLTDFVLLNPLADLVFVCTLPFWVAGTV
EWVFTGCKTLRGMATYVNFVSMLTCTITVDFRIVVVOQAKFNRQAKWKIWOVI
CLLWVSLVSLPOLIYGHVODIDKLIQVHSEISTMVLAVIOMTGLFPLPLTML
CYSGLIKULLHARFQKHKSUKIFLVVAVFLQTLPNLANLQCSISWEYTTTSRK
YAVVTEIAIYFRACINLPYAFVGLKFRKNWKLMDIGICLSHLGVSSQWKSSEDS
KTCASHNVETTSFQL"

ORIGIN

Query Match 98.4%; Score 1231.4; DB 11; Length 1791;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1232; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 15 GCGACACCTTTAGTGAGCACACTTCCTGGAACAAAGCTACTGGGCTTCTCTTGAT 74
DB 10 GCGACACCTTTAGTGAGCACACTTCCTGGAACAAAGCTACTGGGCTTCTCTTGAT 69
QY 75 GCGATGGATGATGGGATCAAGAGTCAGTCTGTACGATGGGCACTAGGAGGAGATTTC 134
DB 70 GCGATGGATGATGGGATCAAGAGTCAGTCTGTACGATGGGCACTAGGAGGAGATTTC 129
QY 135 TGGCTCTTCAACAAATCCAGTGATAACAGCCAGGAGAAACAAACGCTTCTTAAAGTCAAG 194
DB 130 TGGCTCTTCAACAAATCCAGTGATAACAGCCAGGAGAAACAAACGCTTCTTAAAGTCAAG 189
QY 195 GAGGCTTTTGGCCCTGTGTACTGTAGTGTGTCTTGGACGTCTAGGAAACTCC 254
DB 190 GAGGCTTTTGGCCCTGTGTACTGTAGTGTGTCTTGGACGTCTAGGAAACTCC 249
QY 255 CTGGTCTTGATTATATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTCTTCTG 314
DB 250 CTGGTCTTGATTATATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTCTTCTG 309
QY 315 CTGAACCTTGGCCCTGGCTGACCTGGTGTCTGTACTCTGCCCTTTTGGGCTATGCA 374
DB 310 CTGAACCTTGGCCCTGGCTGACCTGGTGTCTGTACTCTGCCCTTTTGGGCTATGCA 369
QY 375 GGCACCTATGATGGTCTTTGGCACAGTCAATGCGAAACCTCTCGAGGCACTATACA 434
DB 370 GGCACCTATGATGGTCTTTGGCACAGTCAATGCGAAACCTCTCGAGGCACTATACA 429
QY 435 ATGAACCTTCTACGTCTCAATCTCACTCACTCACTCACTCACTCACTCACTCACTCA 494
DB 430 ATGAACCTTCTACGTCTCAATCTCACTCACTCACTCACTCACTCACTCACTCACTCA 489
QY 495 GTGGTCCAGGCTACCAAGGCTTCAACGGCAGGCTAAGTGAAGATCTGGGCAAGTC 554
DB 490 GTGGTCCAGGCTACCAAGGCTTCAACGGCAGGCTAAGTGAAGATCTGGGCAAGTC 549
QY 555 ATTTGCTTGTCTCATTTGGGTGCTCTCCCTGTGTGTCTTTTGGCACAGATCATCTATGGC 614
DB 550 ATTTGCTTGTCTCATTTGGGTGCTCTCCCTGTGTGTCTTTTGGCACAGATCATCTATGGC 609
QY 615 CATGTTCAAGATATGACAGCTTATCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 674
DB 610 CATGTTCAAGATATGACAGCTTATCTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 669
QY 675 GTTCTTGTATACAGATGACTCTGGGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 734
DB 670 GTTCTTGTATACAGATGACTCTGGGTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 729
QY 735 TACTCAGGCAATATCAAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 794
DB 730 TACTCAGGCAATATCAAGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
QY 795 AAGATCATCTTCTTGTAGT 854
DB 790 AAGATCATCTTCTTGTAGT 849
QY 855 ATGTTAATCCAAAGTCAAGCTGGGAGTACTATACCAATACCAAGCTTTAAGTATGCAATC 914
DB 850 ATGTTAATCCAAAGTCAAGCTGGGAGTACTATACCAATACCAAGCTTTAAGTATGCAATC 909

```
QY 915 GTAGTGAAGAGGCTATAGCATACCTTCGGGCTTGCCTTAACCCCTGACTTATGCTTT 974
Db 910 GTAGTGAAGAGGCTATAGCATACCTTCGGGCTTGCCTTAACCCCTGACTTATGCTTT 969
QY 975 GTTGGCTTAAAGTTCGGGAAGACGCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCT 1034
Db 970 GTTGGCTTAAAGTTCGGGAAGACGCTCTGGAACCTTATGAAGGATATCGGCTGCTCTCT 1029
QY 1035 CACCTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCTCAAGACTTGTTCGCC 1094
Db 1030 CACCTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCTCAAGACTTGTTCGCC 1089
QY 1095 TCCCAATATGAGACACACAGATGTTCCAAATGTAGTAGGCTTGGCCACACTTAGAG 1154
Db 1090 TCCCAATATGAGACACACAGATGTTCCAAATGTAGTAGGCTTGGCCACACTTAGAG 1149
QY 1155 AAGTTAATAACAGAAATCTAGGACATGGCTGTATCATTTGGATGCAACAAGAAAGCTT 1214
Db 1150 AAGTTAATAACAGAAATCTAGGACATGGCTGTATCATTTGGATGCAACAAGAAAGCTT 1209
QY 1215 TGCATTATAGCATGTGGAGTATCATGGAAGCC 1247
Db 1210 TGCATTATAGCATGTGGAGTATCATGGAAGTC 1242

RESULT 2
AY399296 1029 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus CXCR6 gene, VIRTUAL TRANSCRIPT, partial sequence,
ACCESSION AY399296
VERSION AY399296.1 GI:39755285
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1029)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1029)
AUTHORS Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..1029
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1029
/gene="CXCR6"
/locus_tag="HCM0176"

gene
Query Match 82.3%; Score 1029; DB 29; Length 1029;
Best Local Similarity 100.0%; Pred. No. 7.1e-293;
Matches 1029; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
QY 105 CTGTAGGATGGGCACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATACAGC 164
Db 1 CTGTAGGATGGGCACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATACAGC 60
```

```
QY 165 CAGAGAAACAAGCTTCTAAAGTTCAAGAGAGTCTTTTTCGCCCTGTGTACCTGGTA 224
Db 61 CAGAGAAACAAGCTTCTAAAGTTCAAGAGAGTCTTTTTCGCCCTGTGTACCTGGTA 120
QY 225 GTGTTTGTCTTTGAGTGTAGGAACTCCCTGGTCTCTGATTATATACATTTTCTACCAAG 284
Db 121 GTGTTTGTCTTTGAGTGTAGGAACTCCCTGGTCTCTGATTATATACATTTTCTACCAAG 180
QY 285 AAGCTGAGGACTCTGACAGATGTGTTTCTGCTGAACCTTGCCTTGGCTGACCTGGTGT 344
Db 181 AAGCTGAGGACTCTGACAGATGTGTTTCTGCTGAACCTTGCCTTGGCTGACCTGGTGT 240
QY 345 GTCTGTACTCTGCCCTTTTGGGCTATGAGGACCACTATGAGTGGTCTTTTGGSCACAGTC 404
Db 241 GTCTGTACTCTGCCCTTTTGGGCTATGAGGACCACTATGAGTGGTCTTTTGGSCACAGTC 300
QY 405 ATGTGCAAAACTCTTCGAGGCATGTATACAATGAACCTTCTAGCTGCCATGTCTCTC 454
Db 301 ATGTGCAAAACTCTTCGAGGCATGTATACAATGAACCTTCTAGCTGCCATGTCTCTC 360
QY 465 ACCTGCATCAGTGGATGCTTTCATGTAGTGTCCAGGCTACCAAGGCTTCAACCCG 524
Db 361 ACCTGCATCAGTGGATGCTTTCATGTAGTGTCCAGGCTACCAAGGCTTCAACCCG 420
QY 525 CAGGCTAAGTGAAGATCTGGGGCCAAAGTCATTTGCTTGTCTCATTTGGTGGTCTCCCTG 584
Db 421 CAGGCTAAGTGAAGATCTGGGGCCAAAGTCATTTGCTTGTCTCATTTGGTGGTCTCCCTG 480
QY 585 TTGGTTTCTTTGGCCACAGATCATCTATGGCCATGTTCAGATATTCAGAGCTTATCTGT 644
Db 481 TTGGTTTCTTTGGCCACAGATCATCTATGGCCATGTTCAGATATTCAGAGCTTATCTGT 540
QY 645 CAGTACCACAGTCAGGAGATATCCACTATGTTTCTTGTATACAGATGACTCTGGGGTTC 704
Db 541 CAGTACCACAGTCAGGAGATATCCACTATGTTTCTTGTATACAGATGACTCTGGGGTTC 600
QY 705 TTCTGCAATGTCTACTATGATTCGTGCTACTCAGGCAATATCAAGACTTCTTTCAT 764
Db 601 TTCTGCAATGTCTACTATGATTCGTGCTACTCAGGCAATATCAAGACTTCTTTCAT 660
QY 765 GCTCGAAGCTTCAGAGACAAATCTTAAAGATCATCTTCCCTGTAGTGGCTGTGTTTC 824
Db 661 GCTCGAAGCTTCAGAGACAAATCTTAAAGATCATCTTCCCTGTAGTGGCTGTGTTTC 720
QY 825 CTGCTGACCCAGACACCCCTTCAACCTTGGCATGTAAATCCAAAGTACAAAGCTGGAGTAC 884
Db 721 CTGCTGACCCAGACACCCCTTCAACCTTGGCATGTAAATCCAAAGTACAAAGCTGGAGTAC 780
QY 885 TATACCAATACAGCTTAAAGTATGCCATCGTAGTCAGAGGCTATAGCATACTTTCGG 944
Db 781 TATACCAATACAGCTTAAAGTATGCCATCGTAGTCAGAGGCTATAGCATACTTTCGG 840
QY 945 GCTTGCCTTAAACCTGTACTTATGCTTGTGGCTTAAAGTTCCGGAAGAACGCTCGG 1004
Db 841 GCTTGCCTTAAACCTGTACTTATGCTTGTGGCTTAAAGTTCCGGAAGAACGCTCGG 900
QY 1005 AAACCTTAAAGGATATCGGCTGCTCTCTCACTGGAGTCTCAAGTCAATGGAAGTCT 1064
Db 901 AAACCTTAAAGGATATCGGCTGCTCTCTCACTGGAGTCTCAAGTCAATGGAAGTCT 960
QY 1065 TCTGAGGACAGTTCAGAGACTTGTTCCTGCCCTCCACAAATGTAGAGACCAACCATGTGTTTC 1124
Db 961 TCTGAGGACAGTTCAGAGACTTGTTCCTGCCCTCCACAAATGTAGAGACCAACCATGTGTTTC 1020
QY 1125 CAATTGTAG 1133
Db 1021 CAATTGTAG 1029

BY746127 991 bp mRNA linear EST 17-DEC-2002
LOCUS RESULT 3
BY746127
```

DEFINITION

BY746127 RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD) Mus musculus cDNA clone E43004020 5', mRNA sequence.

ACCESSION

BY746127

VERSION

BY746127.1 GI:27173642

KEYWORDS

EST.
Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 (bases 1 to 931)

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H., Iagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Balt, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Bruscia, V., Chotha, C., Corbani, L. E., Cousine, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. D., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempile, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

Nature 420, 563-573 (2002)

MEDLINE

22354683

PUBMED

12466851

COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gs.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken.

Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/WRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

1..991
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="taxon:10090"
/clone="E43004020"
/tissue_type="thymus"
/cell_type="thymic cells"
/clone_lib="RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD)"

ORIGIN

Query Match 68.7%; Score 859; DB 13; Length 991;
Best Local Similarity 95.1%; Pred. No. 1.6e-242;
Matches 938; Conservative 0; Mismatches 43; Indels 5; Gaps 5;
QY 11 CGAGGGCAGACCTTTAGTGAGCACACTTCACTCTGGACAAAGCTACTGGGCTTCTCTTC 70
DB 7 CGAGGGCAGACCTTTAGTGAGCACACTTCACTCTGGACAAAGCTACTGGGCTTCTCTTC 66
QY 71 TGATGCCATGATGATGGGCATCAAGAGTCAGCTCTGTACGATGGCAGCTACGAGGAGA 130
DB 67 TGATGCCATGATGATGGGCATCAAGAGTCAGCTCTGTACGATGGCAGCTACGAGGAGA 126
QY 131 TTTCTGGCTCTTCAACAAATTCAGTGAATACAGCCAGAGAGAACAAAGCTCTCTAAAGTT 190
DB 127 TTTCTGGCTCTTCAACAAATTCAGTGAATACAGCCAGAGAGAACAAAGCTCTCTAAAGTT 186
QY 191 CAAGAGAGTCTTTTGGCCCTGTGTACTCTGTAGTGTGTCTTTGAGCTGTAGGAAA 250
DB 187 CAAGAGAGTCTTTTGGCCCTGTGTACTCTGTAGTGTGTCTTTGAGCTGTAGGAAA 246
QY 251 CTCCTGGTTCTGATTATATATATTTTACCAAGCTGAGGACTCTGACAGATGTGT 310
DB 247 CTCCTGGTTCTGATTATATATATTTTACCAAGCTGAGGACTCTGACAGATGTGT 306
QY 311 TCTGTGAACCTGGCCCTGGCTGACCTGTGTGTCTGTACTCTGCGCTTTTGGGCTA 370
DB 307 TCTGTGAACCTGGCCCTGGCTGACCTGTGTGTCTGTACTCTGCGCTTTTGGGCTA 366
QY 371 TGCAGGCACCTATGAGTGGGTCTTTTGGCACAGTCACTGTGCAAAACTCTTGAGGCATGTA 430
DB 367 TGCAGGCACCTATGAGTGGGTCTTTTGGCACAGTCACTGTGCAAAACTCTTGAGGCATGTA 426
QY 431 TACATGAACTTCTACGTGTCTCACTCTCTCACTGATCACTGATGATGATGATGATGAT 490
DB 427 TACATGAACTTCTACGTGTCTCACTCTCTCACTGATCACTGATGATGATGATGATGAT 486
QY 491 TGTAGTGTCTCCAGGCTACCAAGGCTTCAACCGGCGAGGCTAAAGTGAAGATCTTGGGGCCA 550
DB 487 TGTAGTGTCTCCAGGCTACCAAGGCTTCAACCGGCGAGGCTAAAGTGAAGATCTTGGGGCCA 546
QY 551 AGTCATTGCTTGTCTCACTTGGGTGGTCTCTCTCTTGGTCTTTTGGCCACAGATCATCTA 610
DB 547 AGTCATTGCTTGTCTCACTTGGGTGGTCTCTCTCTTGGTCTTTTGGCCACAGATCATCTA 606
QY 611 TGGCCATCTTCAAGATATTGCAAGCTTATCTGTCACTACCAAGTGAAGAGATATCCAC 670
DB 607 TGGCCATCTTCAAGATATTGCAAGCTTATCTGTCACTACCAAGTGAAGAGATATCCAC 666
QY 671 TATGGTCTTGTATACAGATGACTCTGGGTCTTCTCTGCGCATTTGCTCATGATTTCT 730
DB 667 TATGGTCTTGTATACAGATGACTCTGGGTCTTCTCTGCGCATTTGCTCATGATTTCT 726

QY 731 GTGCTACTCAGGCAATATCAAGACCTTGCTTCATGCTCGAAATCTCCAGAGCAAAATC 790
 Db 727 GTGCAACTCAGGCAATATCAAGACCTTGCTTCATGCTCGAAATCTCCAGAGCAAAATC 786
 QY 791 TCTAAAGATCATCTCTCTGTAGTGGCTGTCTCTGCTGACCCAGACACCCCTTCAACCT 850
 Db 787 TCTAAAGATCATCTCTCTGTAGTGGCTGTCTCTGCTGACCCAGACACCCCTTCAACTA 845
 QY 851 TGCCATGTAATCAAAAGTCAAGCT-CGGAGTACTATACCATTAACAGAGCTTTAAGTATG 909
 Db 846 TGCCATGTAATCCACAGCTCAACGCTGGGAGTACTATACCATGCTTTAATGATG 905
 QY 910 CCATCGTAGTACAGAGGCT-ATAGCATACTTTGGGCTTCCCTTAAC-CCTGTACTTTA 967
 Db 906 CATCCTAGTACAGAGGCTAATAGCATATATCTCTGGCTGCGCTTAANACCTGTACTATA 965
 QY 968 -TGCCTTTGTGGCTTAAAGTTCGG 992
 Db 966 TTGCCTTTGTGGCTTATAGTCCGG 991

RESULT 4
 LOCUS AY399294 1029 bp DNA linear GSS 15-DEC-2003
 DEFINITION Homo sapiens CXCR6 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.

ACCESSION AY399294
 VERSION AY399294.1 GI:39755283

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 1029)
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous

JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302

REFERENCE 2 (bases 1 to 1029)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES Location/Qualifiers

source
 1..1029
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>1029
 /gene="CXCR6"
 /locus_tag="HCM0176"

ORIGIN

Query Match 54.3%; Score 679.2; DB 29; Length 1029;
 Best Local Similarity 81.2%; Pred. No. 2.9e-189;
 Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 166 AGGAGCAAAAGCTTCCTAAAGTTCAGGAGTCTTTTGGCCCTGTGTACCTGGTAG 225

Db 62 AGGAGCATCAAGACTTCTCGAGTTCAGCAAGGCTTTTGGCCCTGCATGACCTGGTGG 121

QY 226 TGTGTGCTTTGGACTGCTAGGAACTCCCTGGTCTGATTATATACATTTTCTACACAGA 285

Db 122 TGTGTGCTGTGGTCTGGTGGGAACTCTCTGGTCTGGTGCATATCCATCTTCAACATA 181

QY 286 AGCTGAGGACTCTGACAGATGTGTTTCTGCTGAACCTTGCCCTGCTGACCTGGTGTGTTG 345
 Db 182 AGTTGCAGAGCCTGACGGATGTGTTCTGCTGAACCTACCTACCTGGCTGACCTGGTGTGTTG 241
 QY 346 TCTGTACTCTGCCCTTTTGGGCTATGACGACCACTTATGAGTGGGTCTTTGGCACAGTCA 405
 Db 242 TCTGCACTCTGCCCTTCTGGGCTATGACGAGCATCCATGAATGGGTGTTTGGCCAGTCA 301
 QY 406 TGTGCAAAAACCTTCTCAGGAGCATGTATACAAATGAACCTTCTACGTGTCCATGCTCACTCTCA 465
 Db 302 TGTGCAAGAGCCTACTGGGCATCTACACTATTAACTTCTACACGTCATGCTCATCTCA 361
 QY 466 CTGTGCATCAAGTGGATCGTTTCATTGTAGTGTCCAGGCTACCAAGGCTTCAACCGGC 525
 Db 362 CTGTGCATCACTGTGGATCGTTTCATTGTAGTGTAAAGGCCCAAGGCTCAACACAGC 421
 QY 526 AGGCTAAGTGGAGATCTGGGGCAAGTCAATTTGCTTGTGCTCAATTTGGGTGGTCTCCCTGT 585
 Db 422 AAGCCAAGAGGATGACCTGGGCAAGTCAACAGCTTGTCTCATCTCGGTGATATCCCTGC 481
 QY 586 TGTGTTCTTTGGCACAGATCACTATGGCCATGTTCAGATATTTGACAAAGCTTATCTGTC 645
 Db 482 TGTGTTCTTTGGCCCAAAATATCTATGGCAATGTCTTTAATCTCGACAAAGCTCATATGTG 541
 QY 646 AGTACCACAGTGAAGGATATCCACTATGTTCTTGTATACAGATGACTCTGGGGTCTTCT 705
 Db 542 GTTACCATGACGAGGCAATTTCCACTGTGGTCTTGGCCACCAGATGACACTGGGGTCTT 601
 QY 706 TCCTGCAATGTCTACTATGATTTCTGTCTACTCAGGCAATATCAAGACTTGTCTTCAATG 765
 Db 602 TCTTGGCACTGTCTCACCATGATTTGTCTGCTATTTCAGTCAATAATCAAAACACTGCTTCA 661
 QY 766 CTGGAACCTTCCAGAGGACAAATCTCTAAAGATCATCTTCTTGTAGTGGTGTGTTCC 825
 Db 662 CTGAGGCTTCCAGAGGACAGATCTCTAAAGATCATCTTCTTGTAGTGGTGTGTTCC 721
 QY 826 TGTGACCCACAGACACCCCTTCAACCTTGGCATGTGTTAATCAAAAGTACAAGCTGGGAGTACT 885
 Db 722 TGTGACCCAGATGCCCTTCAACCTCATGAAGTTCATCCGACGACACACTGGGAATACT 781
 QY 886 ATACCATACAGCTTTAAGTATGCCATCTGTAGTGACAGAGGCTATAGCATCTTTCGGG 945
 Db 782 ATGCCATGACCACTTTCACTACACCATCATGTGTGACAGAGGCCATCGCATCTCAGGG 841
 QY 946 CTTGCCCTTAACCTGTACTTTATGCTTGTGGCTTAAAGTTCGGGAAGAACTCTGGA 1005
 Db 842 CTTGCCCTTAACCTGTCTCTATGCTTGTGACCTGAAGTTCGAAAGAACTTCTGGA 901
 QY 1006 AACTTATGAAGGATATGGCTGCTCTCTCACTGGAGTCTCAAGTCAATGGAAGTCTT 1065
 Db 902 AACTTGTGAAGGACATTTGTTGCTCTCCCTTACCTTGGGGTCTCACAATCAATGGAATCTT 961
 QY 1066 CTGAGGACAGTTCAGAGACTTGTTCGCTCCACAAATGTAGAGACCACAGATATGTTCC 1125
 Db 962 CTGAGGCAATTCAGAGACTTTTTCGCTCCCAATGTGAGGGCCACCAAGATGTTCC 1021
 QY 1126 AATTGTAG 1133
 Db 1022 AGTTATAG 1029

RESULT 5

AY399295

LOCUS

DEFINITION

AY399295

1029 bp DNA linear

GSS 15-DEC-2003

genomic survey sequence.

ACCESSION

AY399295

VERSION

AY399295.1

GI:39755284

GSS.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1 (Bases 1 to 1029)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 14671302
 2 (Bases 1 to 1029)
 Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source Location/Qualifiers
 1..1029
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>1029
 /gene="CXCR6"
 /locus_tag="HCM0176"
 gene
 14671302
 ORIGIN
 Query Match 54.3%; Score 679.2; DB 29; Length 1029;
 Best Local Similarity 80.8%; Pred. No. 2.9e-189;
 Matches 805; Conservative 0; Mismatches 188; Indels 3; Gaps 1;
 141 TTCAACAATTCAGTATACAGCC---AGGAGAACAAAGCTTCTTAAGTTCAGGAG 197
 Db 34 TTCAACAGTTTCAATGACAGCAGCAGGAGGATCATCAAGACTTCTGCGAGTTCAGCAAG 93
 Qy 198 GTCTTTTGGCCCTGTGTGACCTGGTGTGTTCTTTGGAGTCTAGCAAACTCCCTG 257
 Db 94 GTCTTTTGGCCCTGTGTGACCTGGTGTGTTCTTTGGAGTCTAGCAAACTCCCTG 153
 Qy 258 GTTCTGATATATACATTTTCTACCAAGAGCTGAGGACTCTGACAGATGTTTCTGCTG 317
 Db 154 GTGCTGCTATATCCATCTTCTACCAAGTTCAGAGCCTGACGGATGTGTTCTCTGGTG 213
 Qy 318 AACTTGCCCTGCTGACCTGGTGTGTTCTGCTGACTCTGCTGCTTTGGCCCTATGCGAGC 377
 Db 214 AACCTACCTGCTGACCTGGTGTGTTCTGCTGACTCTGCTGCTTTGGCCCTATGCGAGC 273
 Qy 378 ACCTATGAGTGGTCTTTGGCAGAGCTCATGTGCAAACTCTTTCAGGAGTATGATCAATG 437
 Db 274 ATCCATGAATGGTGTGTTGGCCAGGTCTATGTGCAAGAGCTCTAGGCGATCTACACTATT 333
 Qy 438 AACTTCTACGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 497
 Db 334 AACTTCTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 393
 Qy 498 GTCCAGGCTACCAAGGCTTCAACCGGCGAGGCTTAAGTGAAGATCTGGGCGCAAGTCATT 557
 Db 394 GTTAAGGCCCAAGGCTTCAACCGGCGAGGCTTAAGTGAAGATCTGGGCGCAAGTCATT 453
 Qy 558 TGTCTGCTATTTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
 Db 454 AGCTTGCTATCTGGGAGTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 513
 Qy 618 GTTCAAGATATTGACAGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 677
 Db 514 GTCTTTAATCTCCAGGCTCATATGTTGTTACCATGACGAGGCAATTTCCATGTTGTT 573
 Qy 678 CTTGTTATACAGATGACTCTGGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 737
 Db 574 CTTGCCACCAGATGACACTGGGGTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTAT 633

Qy 738 TCAGCATTTATCAAGACCTTGTCTTCTATCTCGAACTTCCAGAGCAAAATCTCTAAAG 797
 Db 634 TCAGTCATAATCAAAACACTGTCTTCTATCTCGAGGCTTCCAGAGCAAGATCTCTAAAG 693
 Qy 798 ATCATCTTCTTGTAGTGGTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 857
 Db 694 ATCATCTTCTTGTAGTGGTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 753
 Qy 858 TTAATCCAAAGTACAAGCTGGAGTACTATACCAATAACCAAGCTTTAAGTATGCCATCGTA 917
 Db 754 CTCATCCGACGACACACTGGGAATACTATGCCATGACCAAGCTTTCATCAACCATCATG 813
 Qy 918 GTGACAGAGCTATAGCATATCTTTCGGGCTTTCCTTAACCTGCTGCTGCTGCTGCTGCTGCTG 977
 Db 814 GTGACAGAGGCTATGCTATCTGAGGCTTTCCTTAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 873
 Qy 978 GCTTTAAAGTTCGGAAGAAGCTTCTGGAACCTTATGAAGGATATCGCTGCTGCTGCTGCTG 1037
 Db 874 AGCTGAGTTCGGAAGAAGCTTCTGGAACCTTATGAAGGATATCGCTGCTGCTGCTGCTGCTG 933
 Qy 1038 CTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCTCAAGAGTCTGCTGCTGCTGCT 1097
 Db 934 CTTGGGCTCTCACATCAATGGAAGTCTTCTGAGGACAAATTCCTCAAGAGTCTTCTGCTGCTGCT 993
 Qy 1098 CACAATGTAGAGCACCAGGATGTTCTCAATTTAG 1133
 Db 994 CACAATGTGAGGACCAAGCATGTTCCAGTTATAG 1029
 BY746487 672 bp mRNA linear EST 17-DEC-2002
 BY746487 RIKEN full-length enriched, 2 days neonate thymus thymic
 cells (NOD) Mus musculus cDNA clone E430014H02 5', mRNA sequence.
 BY746487
 BY746487.1 GI:27174192
 EST
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (Bases 1 to 672)
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,
 Nikaide, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,
 Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,
 Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,
 Batalov, S., Beisel, K.W., Blake, J.A., Bradt, D., Bruscia, V.,
 Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
 Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T.,
 Gariboldi, M., Gliss, C., Godzik, A., Gough, J., Grimmond, S.,
 Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
 Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
 Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R.,
 Maltas, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
 Numata, K., Okido, T., Pavan, W.J., Pertea, G., Pesole, G.,
 Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S.,
 Ravelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
 Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tonita, M.,
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
 Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
 Hayatsu, N., Hitozane-Kishikawa, T., Konno, H., Nakamura, M.,
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
 Shinae, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
 Rogers, J., Birney, E. and Hayashizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL


```
/lab host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1+)"
/notes=vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library [Ref.
Genome Res. 11: 1553-1558 (2001)]. [PMID: 11544199]]. Total
RNAs were obtained from Drs. Dennis Taub, Dan Longo
(National Institute on Aging, USA), Jonathan Keller
(National Cancer Institute, USA). Double-stranded cDNAs
were synthesized with an Oligo(dT) primer (Invitrogen:
5'-pGACTGTTCTAGTACGAGCGCCCTTTTCTTTT-3') from
1.1 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.2 kb. The library was
constructed by Yulan Piao."
```

ORIGIN

Query Match 52.6%; Score 658.2; DB 14; Length 663;
Best Local Similarity 99.5%; Pred. No. 3.6e-183;
Matches 660; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 342 TTTCGTCTGCTACTCGCCCTTTGGGCTATGTCAGGACCTATGAGTGGGTCTTTGGCACA 401
Db 1 TTTCGTCTGCTACTCGCCCTTTGGGCTATGTCAGGACCTATGAGTGGGTCTTTGGGCA 60

Qy 402 GTCATGTGCAAACTCTTCGAGGATGTATACAAATGAACTTCTACGTGTCATGCTCACT 461
Db 61 GTCATGTGCAAACTCTTCGAGGATGTATACAAATGAACTTCTACGTGTCATGCTCACT 120

Qy 462 CTCACCTGCTATCAGTGGATCGTTTCATGTAGTGTCCAGGCTACCAAGGCTTCAAC 521
Db 121 CTCACCTGCTATCAGTGGATCGTTTCATGTAGTGTCCAGGCTACCAAGGCTTCAAC 180

Qy 522 CGGCAGCTAAGTGAAGATCTGGGGCAAGTCATTGTGTCATATTTGGTGGTCTCC 581
Db 181 CGGCAGCTAAGTGAAGATCTGGGGCAAGTCATTGTGTCATATTTGGTGGTCTCC 240

Qy 582 CTGTGTTGTTCTTGGCCACAGATCATCTATGGCCATGTTCAAGATATTCAGAGCTTATC 641
Db 241 CTGTGTTGTTCTTGGCCACAGATCATCTATGGCCATGTTCAAGATATTCAGAGCTTATC 300

Qy 642 TGTCAGTACCAAGTGAAGATATCCATCATGTTCTTGTATTACAGATGACTCTGGGG 701
Db 301 TGTCAGTACCAAGTGAAGATATCCATCATGTTCTTGTATTACAGATGACTCTGGGG 360

Qy 702 TTCTTCCTGCTGCTCATATGATCTGTGCTACTCAGGCAATTCAGAGCTTGTCT 761
Db 361 TTCTTCCTGCTGCTCATATGATCTGTGCTACTCAGGCAATTCAGAGCTTGTCTC 420

Qy 762 CATGCTCGAAATCTCCAGAGCAAAATCTCTAAAGATCATCTCTCTGTAGTGGCTGTG 821
Db 421 CATGCTCGAAATCTCCAGAGCAAAATCTCTAAAGATCATCTCTCTGTAGTGGCTGTG 480

Qy 822 TTCTGCTGACCCAGACACCTTCAACCTTGCCATGTTTAAATCCAAAGTACAGCTGGAG 881
Db 481 TTCTGCTGACCCAGACACCTTCAACCTTGCCATGTTTAAATCCAAAGTACAGCTGGAG 540

Qy 882 TACTATACCAATACCACTTTAAGTATGCGATCGTAGTGACAGAGCTATAGCATCTTT 941
Db 541 TACTATACCAATACCACTTTAAGTATGCGATCGTAGTGACAGAGCTATAGCATCTTT 600

Qy 942 CGGCTTGCCTTAACCTGTACTTATGCTTGTGGCTTAAAGTTCGGAAGACGTC 1001
Db 601 CGGCTTGCCTTAACCTGTACTTATGCTTGTGGCTTAAAGTTCGGAAGACGTC 660

Qy 1002 TGG 1004
Db 661 TGG 663

RESULT 8
CF911440
LOCUS
DEFINITION
A0615D03-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0615D03
IMAGE:30749030 5', mRNA sequence.

ACCESSION
CF911440
VERSION
CF911440.1 GI:38182644
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1 (bases 1 to 647)
Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
Construction of long-transcript enriched cDNA libraries from
submicrogram amounts of total RNAs by a universal PCR amplification
method
Genome Res. 11 (9), 1553-1558 (2001)

JOURNAL
MEDLINE
21429098
PUBMED
11544199
COMMENT
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: A0615 row: D column: 03
Seq primer: M13 Reverse
High quality sequence stop: 647
POLYA=No.

FEATURES
Location/Qualifiers
1..647
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCR"
/db_xref="taxon:10090"
/clone="NIA:A0615D03 IMAGE:30749030"
/dev_stage="Age -10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1+)" cDNA Library (Long 1)"
/note=vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
Site_2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library [Ref.
Genome Res. 11: 1553-1558 (2001)]. [PMID: 11544199]]. Total
RNAs were obtained from Drs. Dennis Taub, Dan Longo
(National Institute on Aging, USA), Jonathan Keller
(National Cancer Institute, USA). Double-stranded cDNAs
were synthesized with an Oligo(dT) primer (Invitrogen:
5'-pGACTGTTCTAGTACGAGCGCCCTTTTCTTTT-3') from
1.1 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The

average insert size is about 2.2 kb. The library was constructed by Yulan Piao."

ORIGIN

Query Match 51.7%; Score 647; DB 14; Length 647;
Best Local Similarity 100.0%; Pred. No. 7.3e-180;
Matches 647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 ATGGGCACTACGAGGAGATTTCTGGCTCTCAACAATTCAGTGATACAGCCAGGAGA 171
Db 1 ATGGGCACTACGAGGAGATTTCTGGCTCTCAACAATTCAGTGATACAGCCAGGAGA 60

QY 172 ACAACGCTTCTCAAGTTCAGGAGGCTTTTGGCCCTGTGTGTACCTGGTGTAGTGTG 231
Db 61 ACAACGCTTCTCAAGTTCAGGAGGCTTTTGGCCCTGTGTGTACCTGGTGTAGTGTG 120

QY 232 TCTTTGGACTGTAGAACTCCCTGGTCTGATATATACATTTCTACCAAGAGCTGA 291
Db 121 TCTTTGGACTGTAGAACTCCCTGGTCTGATATATACATTTCTACCAAGAGCTGA 180

QY 292 GGACTCTGACAGATGTCTTCTGCTGAACCTTGGCCCTGGCTGACCTGGTGTGTCTGTA 351
Db 181 GGACTCTGACAGATGTCTTCTGCTGAACCTTGGCCCTGGCTGACCTGGTGTGTCTGTA 240

QY 352 CTCTGCCCTTTTGGGCTATGAGGACACTATGAGTGGGTCTTTGGCAAGTCATGTGCA 411
Db 241 CTCTGCCCTTTTGGGCTATGAGGACACTATGAGTGGGTCTTTGGCAAGTCATGTGCA 300

QY 412 AAACCTCTCGAGGATGATACAAATGCTTCTAGCTGTCATGCTCACTCTCACTTGTCA 471
Db 301 AAACCTCTCGAGGATGATACAAATGCTTCTAGCTGTCATGCTCACTCTCACTTGTCA 360

QY 472 TCACAGTGGATCGTTTCAATGAGTGGTCCAGGCTACCAAGGCTTCAACCGGAGGCTA 531
Db 361 TCACAGTGGATCGTTTCAATGAGTGGTCCAGGCTACCAAGGCTTCAACCGGAGGCTA 420

QY 532 AGTGAAGATCTGGGGCAAGTCATGCTGCTCATTTGGGTGGTCTCCCTGTGTGTTT 591
Db 421 AGTGAAGATCTGGGGCAAGTCATGCTGCTCATTTGGGTGGTCTCCCTGTGTGTTT 480

QY 592 CTTTCCCAAGATCATCTATGGCCATGTTCAGATATTCAGAACTTATCTGTCAAGTACC 651
Db 481 CTTTCCCAAGATCATCTATGGCCATGTTCAGATATTCAGAACTTATCTGTCAAGTACC 540

QY 652 ACAGTGAGGAGATATCCATATGTTCTGTTATACAGATGACCTGGGGTCTTCTCTGC 711
Db 541 ACAGTGAGGAGATATCCATATGTTCTGTTATACAGATGACCTGGGGTCTTCTCTGC 600

QY 712 CATTCCTCACTATGATTTCTGTCTACTCAGGCATTTATCAAGACCTTG 758
Db 601 CATTCCTCACTATGATTTCTGTCTACTCAGGCATTTATCAAGACCTTG 647

RESULT 9
CF913009
LOCUS 631 bp mRNA linear EST 05-NOV-2003
DEFINITION A0641F09-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+) CDNA Library (long 1) Mus musculus CDNA clone NIA:A0641F09
IMAGE:30751556 5', mRNA sequence.
CF913009
ACCESSION CF913009.1 GI:38184213
VERSION EST.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclauognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 631)
AUTHORS Piao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
TITLE Construction of long-transcript enriched cDNA libraries from submicrogram amounts of total RNAs by a universal PCR amplification method
JOURNAL Genome Res. 11 (9), 1553-1558 (2001)
MEDLINE 21429098

PUBMED COMMENT

11544199
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: A0641 row: F column: 09
Seq primer: M13 Reverse
High quality sequence stop: 631
POLYA=No.

Location/Qualifiers
1. 631
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCr"
/db_xref="nia:EST:A0641F09-5"
/db_xref="taxon:10090"
/clone="NIA:A0641F09 IMAGE:30751556"
/dev_stage="Age -10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1+) CDNA Library (long 1)"
/note="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library [Ref.
Genome Res. 11: 1553-1558 (2001). PMID: 11544199]. Total
RNAs were obtained from Drs. Dennis Taub, Dan Longo
(National Institute on Aging, USA), Jonathan Keller
(National Cancer Institute, USA). Double-stranded cDNAs
were synthesized with an oligo(dT) primer [Invitrogen:
5'-pGACTAGTCTAGTCGAGCGCGCCCTTTTCTTTT-3'] from
1.1 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Loner-linker LL-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricion
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.2 kb. The library was
constructed by Yulan Piao."

ORIGIN

Query Match 50.4%; Score 631; DB 14; Length 631;
Best Local Similarity 100.0%; Pred. No. 3.9e-175;
Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 112 ATGGGCACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATACAGCCAGGAGA 171
Db 1 ATGGGCACTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATACAGCCAGGAGA 60

QY 172 ACAACGCTTCTCAAGTTCAGGAGGCTTTTGGCCCTGTGTGTACCTGGTGTAGTGTG 231
Db 61 ACAACGCTTCTCAAGTTCAGGAGGCTTTTGGCCCTGTGTGTACCTGGTGTAGTGTG 120

QY 232 TCTTTGGACTGTAGAACTCCCTGGTCTGATATATACATTTTCTACCAAGAGCTGA 291
Db 121 TCTTTGGACTGTAGAACTCCCTGGTCTGATATATACATTTTCTACCAAGAGCTGA 180

QY 292 GGACTCTGACAGATGTCTTCTGCTGAACCTTGGCCCTGGCTGACCTGGTGTGTCTGTA 351
Db 181 GGACTCTGACAGATGTCTTCTGCTGAACCTTGGCCCTGGCTGACCTGGTGTGTCTGTA 240

QY 352 CTCTGCCCTTTTGGGCTATGAGGACACTATGAGTGGGTCTTTGGCAAGTCATGTGCA 411
Db 241 CTCTGCCCTTTTGGGCTATGAGGACACTATGAGTGGGTCTTTGGCAAGTCATGTGCA 300

QY 412 AAACCTCTCGAGGATGATACAAATGCTTCTAGCTGTCATGCTCACTCTCACTTGTCA 471

Db 301 AAATCTTCAGAGGATGATACATGAATGAATCTTACGTTCTCATGCTCACTCTCACTGCA 360
 Qy 472 TCACAGTGGATCGTTTCATTTGATGTGTCAGGTCACCAAGGCTTCAACCGCAGGCTA 531
 Db 361 TCACAGTGGATCGTTTCATTTGATGTGTCAGGTCACCAAGGCTTCAACCGCAGGCTA 420
 Qy 532 AGTGGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 591
 Db 421 AGTGGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 480
 Qy 592 CTTTGGCCACAGATCATCTATGGCCATTTCAAGATATTGAAGCTTATCTGTCAGTACC 651
 Db 481 CTTTGGCCACAGATCATCTATGGCCATTTCAAGATATTGAAGCTTATCTGTCAGTACC 540
 Qy 652 ACAGTGGAGATATCCACTATGTTCTTGTATACAGATGACTCTGGGGTCTTCTCTGC 711
 Db 541 ACAGTGGAGATATCCACTATGTTCTTGTATACAGATGACTCTGGGGTCTTCTCTGC 600
 Qy 712 CATTGCTCACTATGATCTGCTACTCAGG 742
 Db 601 CATTGCTCACTATGATCTGCTACTCAGG 631

RESULT 10
 CF911103
 LOCUS
 DEFINITION A0610D08-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
 cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0610D08
 IMAGE:30748555 5', mRNA sequence.
 CF911103
 VERSION 1
 KEYWORDS
 SOURCE
 ORGANISM Mus musculus (house mouse)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1. (bases 1 to 627)
 Pao, Y., Ko, N.T., Lim, M.K. and Ko, M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)
 21429098
 11544199
 COMMENT
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: A0610 row: D column: 08
 Seq primer: M13 Reverse
 High quality sequence stop: 627
 POLYA=No.

FEATURES
 source
 1. 627
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6NCr"
 /db_xref="niaEST:A0610D08-5"
 /db_xref="taxon:10090"
 /clone="NIA:A0610D08 IMAGE:30748555"
 /dev_stage="Age ~10 weeks old"
 /lab_host="DH10B"
 /clone_lib="NIA Mouse Hematopoietic Stem Cell
 (Lin-/c-Kit-/Sca-1+) cDNA Library (Long 1)"
 /note="vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
 Site_2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).
 This is a long-transcript enriched cDNA library (Ref.
 Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
 RNAs were obtained from Drs. Dennis Taub, Dan Longo

(National Institute on Aging, USA), Jonathan Keller
 (National Cancer Institute, USA). Double-stranded cDNAs
 were synthesized with an Oligo(dT) primer (Invitrogen;
 5'-TGACTATCTTCAGATCGGAGCGGCCCTTTTCTTTT-3') from
 1.1 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to lone-linker LL-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
 vector. The DH10B E. coli host was transformed with the
 ligation mixture by the standard chemical method. The
 average insert size is about 2.2 kb. The library was
 constructed by Yulan Piao."

ORIGIN

Query Match 50.1%; Score 627; DB 14; Length 627;
 Best Local Similarity 100.0%; Pred. No. 6e-174;
 Matches 627; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 112 ATGGGCACCTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATAACAGCCAGGAGA 171
 Db 1 ATGGGCACCTACGAGGAGATTTCTGGCTCTTCAACAATTCAGTGATAACAGCCAGGAGA 60
 Qy 172 ACAACGGCTTCTCTAAAGTTCAAGAGGTCCTTTTGGCCCTGTGTACTCTGGTAGTGTG 231
 Db 61 ACAACGGCTTCTCTAAAGTTCAAGAGGTCCTTTTGGCCCTGTGTACTCTGGTAGTGTG 120
 Qy 232 TCTTTGGACTCTAGGAACCTCCCTGTTCTGATTATATATACATTTCTACCAAGCTGA 291
 Db 121 TCTTTGGACTCTAGGAACCTCCCTGTTCTGATTATATATACATTTCTACCAAGCTGA 180
 Qy 292 GGAATCTTCACAGATGTGTTTCTGCTGAACCTTGCCTGGCTGACCTGGTGTGTTCTGTA 351
 Db 181 GGAATCTTCACAGATGTGTTTCTGCTGAACCTTGCCTGGCTGACCTGGTGTGTTCTGTA 240
 Qy 352 CTCTGCCCTTTTGGCCCTATGACGACCTATAGTGGGCTTTGGCACAGTCATGTGCA 411
 Db 241 CTCTGCCCTTTTGGCCCTATGACGACCTATAGTGGGCTTTTGGCACAGTCATGTGCA 300
 Qy 412 AAATCTTCGAGGATGTATACAACTTCTACGTGTCCATGCTCACTCTCACTGCA 471
 Db 301 AAATCTTCGAGGATGTATACAACTTCTACGTGTCCATGCTCACTCTCACTGCA 360
 Qy 472 TCACAGTGGATCGTTTCATTTGATGTGTCAGGCTACCAAGGCTTCAACCGCAGGCTA 531
 Db 361 TCACAGTGGATCGTTTCATTTGATGTGTCAGGCTACCAAGGCTTCAACCGCAGGCTA 420
 Qy 532 AGTGGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 591
 Db 421 AGTGGAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTCTCCCTCTTGTTT 480
 Qy 592 CTTTGGCCACAGATCATCTATGGCCATTTCAAGATATTGAAGCTTATCTGTCAGTACC 651
 Db 481 CTTTGGCCACAGATCATCTATGGCCATTTCAAGATATTGAAGCTTATCTGTCAGTACC 540
 Qy 652 ACAGTGGAGATATCCACTATGTTCTTGTATACAGATGACTCTGGGGTCTTCTCTGC 711
 Db 541 ACAGTGGAGATATCCACTATGTTCTTGTATACAGATGACTCTGGGGTCTTCTCTGC 600
 Qy 712 CATTGCTCACTATGATCTGCTACT 738
 Db 601 CATTGCTCACTATGATCTGCTACT 627

RESULT 11
 CF912382
 LOCUS
 DEFINITION A0631A05-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
 cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0631A05
 628 bp mRNA linear EST 05-NOV-2003
 CF912382
 LOCUS
 DEFINITION A0631A05-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
 cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0631A05

/db_xref="niaEST:K0616E05-5N"
 /db_xref="taxon:10090"
 /clone="NIA:K0616E05 IMAGE:30070324"
 /tissue_type="Hematopoietic Stem Cell
 (Lin-/c-Kit-/Sca-1+)"
 /dev_stage="Age approx.10 weeks old"
 /lab_host="DRI0B"
 /clone_lib="NIA Mouse Hematopoietic Stem Cell
 (Lin-/c-Kit-/Sca-1+) cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
 NotI; Mouse cDNA project by the Laboratory of Genetics,
 National Institute on Aging (NIA), Intramural Research
 Program, NIH (http://lsun.grc.nia.nih.gov/cDNA). This is
 a long-transcript enriched cDNA library (Ref. Genome Res.
 11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were
 obtained from Drs. Dennis Taub, Dan Longo (National
 Institute on Aging, USA), Jonathan Keller (National Cancer
 Institute, USA). Double-stranded cDNAs were synthesized
 with an Oligo(dT) primer (Invitrogen).
 5'-pGATAGTCTAGATCGAGCGCGCCCTTTTCTTTT-3' from
 1.1 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to lona-linker L1-Sal4, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.
 The DH10B E. coli host was transformed with the ligation
 mixture by the standard chemical method. The average
 insert size is about 2.2 kb. The library was constructed
 by Yulan Piao (NIA)."

Query Match 49.4%; Score 617.4; DB 14; Length 619;
 Best Local Similarity 99.8%; Pred. No. 4.1e-171;
 Matches 618; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 112 ATGGGCACTAGAGGAGATTCTGGCTCTACACATTCAGTGATACAGCCAGGAGA 171
 DB 1 ATGGGCACTAGAGGAGATTCTGGCTCTACACATTCAGTGATACAGCCAGGAGA 60
 QY 172 ACAACCGTCTCTAAAGTTCAAGGAGTCTTTTGGCCCTGTGTGATCCTGGTAGTGTG 231
 DB 61 ACAACCGTCTCTAAAGTTCAAGGAGTCTTTTGGCCCTGTGTGATCCTGGTAGTGTG 120
 QY 232 TCTTTGGACTCTAGGAACCTCCCTGGTCTGATATATACATTTTCTACAGAGCTGA 291
 DB 121 TCTTTGGACTCTAGGAACCTCCCTGGTCTGATATATACATTTTCTACAGAGCTGA 180
 QY 292 GGACTCTACAGAGTGTGTTCTGTGACTTGGCCCTGGCTGACCTGGTGTGTGTTCTGTGA 351
 DB 191 GGACTCTACAGAGTGTGTTCTGTGACTTGGCCCTGGCTGACCTGGTGTGTGTTCTGTGA 240
 QY 352 CTCTGCCCTTTTGGCCCTATCGAGCACTATAGTGGTCTTTGGCAGCATGTGCTGA 411
 DB 241 CTCTGCCCTTTTGGCCCTATCGAGCACTATAGTGGTCTTTGGCAGCATGTGCTGA 300
 QY 412 AAACCTCTCGAGGATGATACATGAATCTTACGTGTCATGCTCCTACCTGCTGA 471
 DB 301 AAACCTCTCGAGGATGATACATGAATCTTACGTGTCATGCTCCTACCTGCTGA 360
 QY 472 TCACAGTGGATCGTTTCAATAGTGGTCCAGGCTACCAAGCCCTTCAACCGGAGGCTA 531
 DB 361 TCACAGTGGATCGTTTCAATAGTGGTCCAGGCTACCAAGCCCTTCAACCGGAGGCTA 420
 QY 532 AGTGAAGATCTGGGCGCAAGTCATTTGCTGCTCAATTTGGTGGTCTCCCTGTGTTT 591
 DB 421 AGTGAAGATCTGGGCGCAAGTCATTTGCTGCTCAATTTGGTGGTCTCCCTGTGTTT 480
 QY 592 CTTTGGCCACAGATCATCTATGGCCATGTTTCAAGATATTGACAAGCTTATCTGTGCTG 651

DB 481 CTTTGCCACAGATCATCTATGGCCATGTTCAAGATATTGACAAGCTTATCTGTGCTG 540
 QY 652 ACAGTGAGGAGATATCCACTATGTTCTTGTATACAGATGACTCTGGGTTCTTCTGTC 711
 DB 541 ACAGTGAGGAGATATCCACTATGTTCTTGTATACAGATGACTCTGGGTTCTTCTGTC 600
 QY 712 CATTGCTCACTATGATTCT 730
 DB 601 CATTGCTCACTATGATTCT 619

BB662193 626 bp mRNA linear EST 26-OCT-2001
 BB662193 RIKEN full-length enriched, 16 days neonate heart Mus
 musculus cDNA clone D830030F09 5', mRNA sequence.
 BB662193
 BB662193.1 GI:16495948
 EST.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 626)
 Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T.,
 Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J.,
 Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K.,
 Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K.,
 Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
 Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
 Takeda,Y., Tanaka,T., Toyota,T., Muramatsu,M. and Hayashizaki,Y.
 RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 Unpublished (2001)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.go.jp
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
 Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
 Watahiki,M., Yoneda,Y., Iehikawa,T., Ozawa,K., Tanaka,T.,
 Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A.,
 and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P.,
 Sugahara,Y. and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I.,
 Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
 Hayashizaki,Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for
 further details.
 e mouse tissues.
 Location/Qualifiers
 1. 626
 /organism="Mus musculus"
 /molecule="mRNA"
 /db_xref="taxon:10090"
 /clone="D830030F09"

FEATURES
 source

ORIGIN		Query Match Best Local Similarity 49.2%; Score 616; DB 10; Length 626; Matches 61; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	15	GGCAGACCTTTAGTGAGCACACTTCACTCTGGAAACAAAGCTACTGGGCTTCTCTTCTGAT	74		
DB	10	GGCAGACCTTTAGTGAGCACACTTCACTCTGGAAACAAAGCTACTGGGCTTCTCTTCTGAT	69		
QY	75	GCCATGATGATGGGCATCAAGTAGTACGTCTGTACGATGGGCACTACGAGGAGGATTC	134		
DB	70	GCCATGATGATGGGCATCAAGTAGTACGTCTGTACGATGGGCACTACGAGGAGGATTC	129		
QY	135	TGCTCTTCAACAATTCAGGTGATPAAACGACGAGAGAAACAAAGCTTCCTAAAGTTCAAG	194		
DB	130	TGGCTCTTCAACAATTCAGGTGATPAAACGACGAGAGAAACAAAGCTTCCTAAAGTTCAAG	189		
QY	195	GAGGTCCTTTTGCCCTGTGTACCTGGTAGTGTGTCTTTGGACTGTAGGAACTCC	254		
DB	190	GAGGTCCTTTTGCCCTGTGTACCTGGTAGTGTGTCTTTGGACTGTAGGAACTCC	249		
QY	255	CTGGTCTGTATTATATACATTTTCTACCAAGACTGAGGACTCTGACAGATGTGTTTCTG	314		
DB	250	CTGGTCTGTATTATATACATTTTCTACCAAGACTGAGGACTCTGACAGATGTGTTTCTG	309		
QY	315	CTGAACCTGGCCCTGGCTGACCTGGTGTGTCTGTACTCTGCCTTTTGGGCGCTATGCA	374		
DB	310	CTGAACCTGGCCCTGGCTGACCTGGTGTGTCTGTACTCTGCCTTTTGGGCGCTATGCA	369		
QY	375	GGCACCTATCAGTGGGCTTTTGGCACAGTCAATGTGCAAAACTCTTCGAGGCATGTATACA	434		
DB	370	GGCACCTATCAGTGGGCTTTTGGCACAGTCAATGTGCAAAACTCTTCGAGGCATGTATACA	429		
QY	435	ATGAACCTTCTACGTGTCCATGCTCACTCTCACCTGCACTACAGTGGATCGTTTCATGTGA	494		
DB	430	ATGAACCTTCTACGTGTCCATGCTCACTCTCACCTGCACTACAGTGGATCGTTTCATGTGA	489		
QY	495	GTGGTCAGGCTACCAAGGCTTCAACCGCAGGCTAAGTGGAGATCTGGGCGCAAGTC	554		
DB	490	GTGGTCAGGCTACCAAGGCTTCAACCGCAGGCTAAGTGGAGATCTGGGCGCAAGTC	549		
QY	555	ATTGTGCTGTCAATTGGGGTGTCTCCCTGTGTGTTTCTTTGGCACAGATCACTATGGC	614		
DB	550	ATTGTGCTGTCAATTGGGGTGTCTCCCTGTGTGTTTCTTTGGCACAGATCACTATGGC	609		
QY	615	CATGTTCAAGATATTG	630		
DB	610	CATGTTCAAGATATTG	625		

RESULT 14
CF910757

CF910757 610 bp mRNA linear EST 05-NOV-2003
 A0604D11-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
 cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0604D11
 IMAGE:30747982 5', mRNA sequence.

CF910757
 CF910757.1 GI:38181933
 Est.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 610)
 Piao,Y., Ko,N.T., Lim,M.K. and Ko,M.S.H.
 Construction of long-transcript enriched cDNA libraries from
 submicrogram amounts of total RNAs by a universal PCR amplification
 method
 Genome Res. 11 (9), 1553-1558 (2001)
 21429098
 11544199
 Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgan.grc.nia.nih.gov
 Plate: A0604 row: D column: 11
 Seq primer: M13 Reverse
 High quality sequence stop: 610
 POLYA=No.

Location/Qualifiers
 1. .610
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="C55EL/6NCr"
 /db_xref="niaEST:A0604D11-5"
 /db_xref="taxon:10090"
 /clone="NIA:A0604D11 IMAGE:30747982"
 /dev_stage="Age ~10 weeks old"
 /lab_host="DH108"
 /clone_lib="NIA Mouse Hematopoietic Stem Cell
 (Lin-/C-Kit-/Sca-1+; cDNA Library (Long 1))"
 /note="Vector: pCMV-SPORT6 (Invitrogen); Site_1: SalI;
 Site_2: NotI; Mouse cDNA project by the Laboratory of
 Genetics, National Institute on Aging (NIA), Intramural
 Research Program, NIH (<http://lgan.grc.nia.nih.gov/cDNA>).
 This is a long-transcript enriched cDNA library (Ref.
 Genome Res. 11: 1553-1558 (2001)). [PMID: 11544199]. Total
 RNAs were obtained from Drs. Dennis Taub, Dan Longo
 (National Institute on Aging, USA), Jonathan Keller
 (National Cancer Institute, USA), Double-stranded cDNAs
 were synthesized with an Oligo(dt) primer [Invitrogen:
 5'-pGACTAGTCTAGATCGAGCGCGCCCTTTTTTT-3'] from
 1.1 ug of total RNA, treated with T4 DNA polymerase, and
 purified by ethanol-precipitation. The cDNAs were ligated
 to Loner-linker LL-SalI, purified by phenol/chloroform, and
 separated from free linkers by Centricon 100. Then, the
 cDNAs were amplified by long-range high fidelity PCR using
 Ex Taq polymerase (Takara) with a primer Sal4-S. The
 products were purified by phenol/chloroform and Centricon
 100. The cDNAs were digested with SalI and NotI enzymes
 and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
 vector. The DH10B E. coli host was transformed with the
 ligation mixture by the standard chemical method. The
 average insert size is about 2.2 kb. The library was
 constructed by Yulan Piao."

ORIGIN

Constructed by Julian Frazar.

Query Match 48.8%; Score 610; DB 14; Length 610;
Best Local Similarity 100.0%; Pred. No. 6.4e-169;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 139 TCTTCAACAATTCAGTGATAACAGCCAGGAGAACAAACGCTTCTCTAAAGTTCGAAGGAGG 138

Db 1 TCTTCAACAATTCAGTGATAACAGCCAGGAGAACAAACGCTTCTCTAAAGTTCGAAGGAGG 60

```
199 TCTTTTGGCCCTGTGTACTGCTAGTGTGTTGTTGTTGACCTGCTAGGAACTCCCTGG 258
Db 61 TCTTTTGGCCCTGTGTACTGCTAGTGTGTTGTTGTTGACCTGCTAGGAACTCCCTGG 120
Qy 259 TTTGATTATATACATTTTTCACAGAGCTGAGGACTCTGACAGATGTTGTTCTGCTGA 318
Db 121 TTTGATTATATACATTTTTCACAGAGCTGAGGACTCTGACAGATGTTGTTCTGCTGA 180
Qy 319 ACTTGCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 378
Db 181 ACTTGCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Qy 379 CCTATGAGTGGGCTTTTGGCACAGTCAATGTCGAAAACCTCTTCGAGGCATGTATACAATGA 438
Db 241 CCTATGAGTGGGCTTTTGGCACAGTCAATGTCGAAAACCTCTTCGAGGCATGTATACAATGA 300
Qy 439 ACTTCTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
Db 301 ACTTCTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Qy 499 TCCAGGCTACCAAGGCTTCAACCGGACGCTAAGTGGAGATCTGGGGCCAACTCATTT 558
Db 361 TCCAGGCTACCAAGGCTTCAACCGGACGCTAAGTGGAGATCTGGGGCCAACTCATTT 420
Qy 559 GCTTGTCTATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 618
Db 421 GCTTGTCTATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 739 CAGGCATTAT 748
Db 601 CAGGCATTAT 610

RESULT 15
CF913351
LOCUS
DEFINITION
  CF913351 601 bp mRNA linear EST 05-NOV-2003
  A0647A08-5 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1+)
  cDNA Library (Long 1) Mus musculus cDNA clone NIA:A0647A08
  IMAGE:30752071 5', mRNA sequence.
ACCESSION
  CF913351
VERSION
  CF913351.1 GI:38184555
KEYWORDS
  EST.
SOURCE
  Mus musculus (house mouse)
ORGANISM
  Mus musculus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
  1 (bases 1 to 601)
  Piao Y., Ko N.T., Lim M.K. and Ko M.S.H.
  Construction of long-transcript enriched cDNA libraries from
  submicrogram amounts of total RNAs by a universal PCR amplification
  method
  Genome Res. 11 (9), 1553-1558 (2001)
JOURNAL
  MEDLINE
  PUBMED
  21429098
  11544199
COMMENT
  Contact: Dawood B. Dudekula
  Laboratory of Genetics
  National Institute on Aging/National Institutes of Health
  333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
  Email: cdna@igsun.grc.nia.nih.gov
  Plate: A0647 row: A column: 08
  Seg primer: M13 Reverse
  High quality sequence stop: 601
  POLYA=No. Location/Qualifiers
```

source

```
1. 601
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6Ncr"
/db_xref="niaEST:A0647A08-5"
/db_xref="taxon:10090"
/clones="NIA:A0647A08 IMAGE:30752071"
/dev_stage="Age ~10 weeks old"
/lab_host="DH10B"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1+) cDNA Library (Long 1)"
/notes="Vector: pCMV-SPORT6 (Invitrogen); Site 1: SalI;
Site 2: NotI; Mouse cDNA project by the Laboratory of
Genetics, National Institute on Aging (NIA), Intramural
Research Program, NIH (http://igsun.grc.nia.nih.gov/cDNA).
This is a long-transcript enriched cDNA library (Ref.
Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total
RNAs were obtained from Drs. Dennis Taub, Dan Longo
(National Institute on Aging, USA), Jonathan Keller
(National Cancer Institute, USA), Double-stranded cDNAs
were synthesized with an Oligo(dT) primer [Invitrogen:
5'-PGATAGTCTAGATCGAGCGCCGCTTTTCTTTTCTTTT-3'] from
1.1 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lene-linker Lb-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pCMV-SPORT6 plasmid
vector. The DH10B E. coli host was transformed with the
ligation mixture by the standard chemical method. The
average insert size is about 2.2 kb. The library was
constructed by Yulan Piao."
```

ORIGIN

```
Query Match 47.9%; Score 599.4; DB 14; Length 601;
Best Local Similarity 99.8%; Pred. No. 8.7e-166;
Matches 600; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 139 TCTTCAACAATTCAGTGTATACAGCCAGGAGAACAAACGCTTCTTAAGTTCAGGAGG 198
Db 1 TCTTCAACAATTCAGTGTATACAGCCAGGAGAACAAACGCTTCTTAAGTTCAGGAGG 60
QY 199 TCTTTTGGCCCTGTGTACTGCTAGTGTGTTGTTGTTGACCTGCTAGGAACTCCCTGG 258
Db 61 TCTTTTGGCCCTGTGTACTGCTAGTGTGTTGTTGTTGACCTGCTAGGAACTCCCTGG 120
QY 259 TTTGATTATATACATTTTTCACAGAGCTGAGGACTCTGACAGATGTTGTTCTGCTGA 318
Db 121 TTTGATTATATACATTTTTCACAGAGCTGAGGACTCTGACAGATGTTGTTCTGCTGA 180
QY 319 ACTTGCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 378
Db 181 ACTTGCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
QY 379 CCTATGAGTGGGCTTTTGGCACAGTCAATGTCGAAAACCTCTTCGAGGCATGTATACAATGA 438
Db 241 CCTATGAGTGGGCTTTTGGCACAGTCAATGTCGAAAACCTCTTCGAGGCATGTATACAATGA 300
QY 439 ACTTCTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 498
Db 301 ACTTCTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
QY 499 TCCAGGCTACCAAGGCTTCAACCGGACGCTAAGTGGAGATCTGGGGCCAACTCATTT 558
Db 361 TCCAGGCTACCAAGGCTTCAACCGGACGCTAAGTGGAGATCTGGGGCCAACTCATTT 420
QY 559 GCTTGTCTATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 618
Db 421 GCTTGTCTATTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
```

QY 619 TTCACAGATATTCACAAAGCTTATCTCTCAGTACCCACAGTGAGGAGATATCCACTATGGTTC 678
Db |||||
481 TTCACAGATATTCACAAAGCTTATCTCTCAGTACCCACAGTGAGGAGATATCCACTATGGTTC 540
QY 679 TTGTTATACAGATGACTCTGGGGTTCTTCCTGCGCCATTGCTCACTATGATTCTGTGCTACT 738
Db |||||
541 TTGTTATACAGATGACTCTGGGGTTCTTCCTGCGCCATTGCTCACTATGATTCTGTGCTACT 600
QY 739 C 739
Db 601 C 601

Search completed: September 29, 2004, 02:59:53
Job time : 3122 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2004, 00:33:29 ; Search time 120 Seconds
(without alignments)
5785.363 Million cell updates/sec

Title: US-10-603-141-1
Perfect score: 1251
Sequence: 1 gaattcgacagggcaga.....gtatcatggagaagcctgac 1251

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	706.4	56.5	2238	1 US-08-742-011-1	Sequence 1, Appli
2	689.4	55.1	1029	3 US-09-116-498-5	Sequence 5, Appli
3	685	54.8	1037	3 US-09-116-498-3	Sequence 3, Appli
4	679.2	54.3	1029	4 US-09-449-437A-1	Sequence 1, Appli
5	676.8	54.1	1026	3 US-09-275-384B-4	Sequence 4, Appli
6	648	51.8	1026	3 US-09-116-498-1	Sequence 1, Appli
7	183.4	14.7	1074	4 US-09-170-496D-23	Sequence 23, Appli
8	183.4	14.7	2577	4 US-09-266-464-1	Sequence 1, Appli
9	183.4	14.7	2577	4 US-09-016-434-1095	Sequence 1095, Ap
10	181.2	14.5	2751	1 US-08-153-848-23	Sequence 23, Appl
11	181.2	14.5	2751	3 US-09-299-843A-23	Sequence 23, Appl
12	181.2	14.5	2751	4 US-09-088-337B-23	Sequence 23, Appl
13	181.2	14.5	2751	5 PCT-US93-11153-23	Sequence 23, Appl
14	178.6	14.3	1074	4 US-09-170-496D-175	Sequence 175, App
15	177.8	14.2	2085	3 US-09-299-843A-65	Sequence 65, Appl
16	177.8	14.2	2085	4 US-09-088-337B-65	Sequence 65, Appl
17	162	12.9	1255	4 US-09-023-655-993	Sequence 993, App
18	162	12.9	1679	4 US-09-016-434-1097	Sequence 1097, Ap
19	162	12.9	2100	4 US-09-495-050A-289	Sequence 289, App
20	161.4	12.9	1137	4 US-09-170-496D-73	Sequence 73, Appl
21	161.4	12.9	1137	4 US-09-170-496D-203	Sequence 203, App
22	161.4	12.9	1500	4 US-09-023-655-1465	Sequence 1455, Ap
23	161.4	12.9	1900	1 US-08-153-848-18	Sequence 18, Appl
24	161.4	12.9	1900	3 US-09-299-843A-18	Sequence 18, Appl
25	161.4	12.9	1900	4 US-09-088-337B-18	Sequence 18, Appl
26	161.4	12.9	1900	5 PCT-US93-11153-18	Sequence 18, Appl
27	161.4	12.9	2058	1 US-08-153-848-6	Sequence 6, Appli

Sequence 6, Appli
Sequence 6, Appli
Sequence 6, Appli
Sequence 14, Appli
Sequence 14, Appli
Sequence 14, Appli
Sequence 14, Appli
Sequence 970, App
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1214, Ap
Sequence 1094, Ap
Sequence 1, Appli
Sequence 7, Appli
Sequence 3, Appli
Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-742-011-1
; Sequence 1, Application US/08742011
; Patent No. 5824504
; GENERAL INFORMATION:
; APPLICANT: Elshourbagy, Nabil A.
; APPLICANT: Bergsma, Derk J.
; APPLICANT: Ellis, Catherine E.
; TITLE OF INVENTION: Human 7-Transmembrane Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Smithkline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,011
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Har, William T.
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50020P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2238 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 332..1357
US-08-742-011-1

Query Match 56.5%; Score 706.4; DB 1; Length 2238;
Best Local Similarity 78.6%; Pred. No. 6.3e-224;
Matches 871; Conservative 0; Mismatches 231; Indels 6; Gaps 2;
141 TTCACAAATCCAGTGATACAGCC---AGGAGACAACGCTTCCTAAAGTTCAAGGAG 197

QY 358 CTTTGTGGCCCTATGACGGCACCTATGAGTGGGCTTTGGCACAGTCAATGTGCAAAATC 417
DB 254 CTTTGTGGCCCTATGACGGCATCATGAATGATCTTTGGCCAGGTCAATGTGCAAGACCC 313
QY 418 TTCAGGCACTATATACATGAATCAATCTTACGTTGTCAGTCTCACTCTCACTGATCAG 477
DB 314 TACTGGGGCTTACACTATTAATCTTACATCATCATCTCACTCTCACTGATCACTG 373
QY 478 TGGATCGTTTCAATGTAGTGTGACAGGTACCAAGGCTTCAACCGGAGCTAAAGTGA 537
DB 374 TGGATCGTTTCAATGTAGTGTGACAGGTACCAAGGCTTCAACCGGAGCTAAAGTGA 433
QY 538 AGATCTGGGGCAAGTCAATCTTGTCTCATTTGGGTGCTCTCTCTGTTGGTTCTTTTC 597
DB 434 TGACTTGGGGCAAGTCAATCTTGTCTCATCTGGGTGATATCCCTGCTGTTCTTCTTGC 493
QY 598 CACAGATCATCTATGGCCATGTTCAAGATATTGACAGCTTATCTGTCAAGTACACAGTG 657
DB 494 CCCAAATATCTATGGCAATGCTTTAATCTGGACAGCTCATATGTTGTTATCATGACA 553
QY 658 AGGAGATATCCACTATGTTCTTGTATACAGATCACTCTGGGTTCTTCTGCTGCAATGC 717
DB 554 AGGAGATTTCCACTGTGTTCTTGCACCCAGATGACACTGGGTTCTTCTTGCACATGC 613
QY 718 TCACATGATCTGTGCTACTCAGGCATATCAAGACCTTGTCTTCACTGCTGAACTTCC 777
DB 614 TCGCCATGATGCTGTCTATTCACTGATCAATCAAAACACTGCTTCACTGAGGCTTCC 673
QY 778 AGAGCACAAATCTCTAAGATCATCTTCTTGTAGTGGCTGTCTCTGCTGACCCAGA 837
DB 674 AGAGCACAGATCTCTAAGATCATCTTCTTGTGATGGCTGTCTCTGCTGACCCAGA 733
QY 838 CACCTTCAACCTTGCCATGTTATTCAAAGTACAGCTGGAGTACTATACATTAACCA 897
DB 734 CACCTTCAACCTTGCCATGTTATTCAAAGTACAGCTGGAGTACTATGCAATGACCA 793
QY 898 GCTTTAGTATGCTGCTAGTACAGAGGTATAGCATCTTTCGGGCTTGGCTTAACC 957
DB 794 GCTTTCACTACCAATCATATGACAGAGGCTATGCAATCTGAGGCTTGGCTTAACC 853
QY 958 CTGTACTTTATGCTTGTGCTTAAAGTTTCGGAAGAGCTCTGGAAGCTTGAAGG 1017
DB 854 CTGTGCTCTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 913
QY 1018 ATATCGGCTGCTCTCTCACTGGAGTCTCAAGTCAAGTCAAGTCTCTGAGGACAGTT 1077
DB 914 ACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 973
QY 1078 CCAAGACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1133
DB 974 CCAAGACTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1029

RESULT 3

US-09-116-498-3
; Sequence 3, Application US/09116498
; Patent No. 6251582

GENERAL INFORMATION:

APPLICANT: Littman, Dan R.
Deng, Hongkui
Unutmaz, Derya
Ramani, Vineet N.K.

TITLE OF INVENTION: NOVEL ALTERNATIVE G-COUPLED RECEPTORS

ASSOCIATED WITH RETROVIRAL ENTRY INTO CELLS, METHODS OF
IDENTIFYING THE SAME, AND DIAGNOSTIC AND THERAPEUTIC USES
THEREOF

NUMBER OF SEQUENCES: 18

CORRESPONDENCE ADDRESS:

ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
Floor
CITY: Hackensack

STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/116,498
FILING DATE: 16-Jul-1998
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-009
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5900
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: African Green Monkey
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-116-498-3

Query Match 54.8%; Score 685; DB 3; Length 1037;

Best Local Similarity 81.1%; Pred. No. 4,8e-217;

Matches 809; Conservative 0; Mismatches 185; Indels 3; Gaps 1;

QY 140 CTTCAACAATTCAGTGTATAACAGCCAGGAGAACAAAC---GCTTCCTAAAGTTCAAGGA 196
DB 33 CTTCAACAGTTTCAATGACAGCAGCCAGGAGGAGCATCAGGACTTCTCGAGTTCAAGCA 92
QY 197 GGTCTTTTGGCCCTGTGTACTGTGTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 256
DB 93 GGTCTTTTGGCCCTGTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 152
QY 257 GGTCTGTATTATATACATTTTCTACAGAAAGCTGAGGACTCTGACAGATGTGTTCGTCT 316
DB 153 GGT 212
QY 317 GAACCTTGGCCCTGGCTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 376
DB 213 GAACCTTGGCCCTGGCTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 272
QY 377 CACCTATGAGTGGTCTTTGGGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 436
DB 273 CATCCATGAATGATCTTTGGCCAGGTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 332
QY 437 GAACCTTGTACGTGTCCATGTCTCACTCTGACATCAGATGAGTGTGTGTGTGTGTGTGTGT 496
DB 333 TAACTTCTACACATCTATGCTCATCTCACTGATCACTGTGTGTGTGTGTGTGTGTGTGTGT 392
QY 497 GGTCCAGGTACCAAGGCTTCAACCGGAGGCTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 556
DB 393 GGTTAAGGCGCCCAAGGCTTATAACCAAGCAAGCAAGTGTGTGTGTGTGTGTGTGTGTGTGT 452
QY 557 TTGCTTGTCTCATTTGGT 616
DB 453 CTGCTTGTCTCATCTGGT 512
QY 617 TGTTCAGATATTGACAAAGCTTATCTGTCAAGTACCACTGAGGAGATATCACTATGTGT 676
DB 513 TGTCTTTAATCTGGCAAGCTCATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 572
QY 577 TCTTGTATTACAGATGACTCTGGGGTCTTCTTCGCCATTGCTCACTATGATTCTGTGTGCTA 736

```
Db 573 TCTTGGCCACCCAGATGACACTGGGGTTCTTCTTCCCACTGCTGCCAATGATTGTCTGCTA 632
Qy 737 CTCAGGCATTAATCAAGACCTTGTCTTCATGCTCGAACTTCCGAAGACACAAATCTCTAAA 796
Db 633 TTCAGTCATAATCAAAACACACTGCTTCATGCTGGAGGCTTCCAGAAGCACAGATCTCTAAA 692
Qy 797 GATCATCTTCTTGTAGTGGGTGTTCTGCTGACCCAGACACCCCTTCAACCTTGGCCAT 856
Db 693 GATCATCTTCTTGTAGTGGGTGTTCTGCTGACCCAGACACCCCTTCAACCTTGGGAA 752
Qy 857 GTTAATCCAAAGTCAAGCTGGGAGTACTATACCATCAACAGCTTTTAAGTATGCCATCGT 916
Db 753 GCTCATCGCAGCACACACTGGGAGTACTATGCCATGACCCAGCTTTCACCTACACCATCAT 812
Qy 917 AGTGACAGAGCTATAGCATACTTTCGGGCTTGGCTTAAACCTGTACTTATAGCCCTTGT 976
Db 813 AGTGACAGAGCCATCGCATACCTGAGGGCTGCTTAAACCTGTACTTATAGCCCTTGT 872
Qy 977 TGGCTTAAAGTTCCGGAAGACGTCGAAACTTATGAAGGATATGCGCTGCTCTCTCA 1036
Db 873 CAGCTGAGTTTCAAGAAACTTCTGGAACCTTGTGAGGACATTTGGCTGTCTCCCTTA 932
Qy 1037 CCTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTTCAAGACTTGTCTGCTC 1096
Db 933 CTTTGGGCTCTCACATCAATGGAATCTTCTGAGGACAAATCCAGACTTTTCTGCTC 992
Qy 1097 CCACAACTAGAGACCCAGTATGTTCCAAATCTAG 1133
Db 993 CCACAACTGGAGGCCACAGCATGTTCCAGTATAG 1029

RESULT 4
US-09-449-437A-1
; Sequence 1, Application US/09449437A
; Patent No. 6319675
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; APPLICANT: Murphy, Kristine E.
; APPLICANT: Wilbanks, Alyson M.
; APPLICANT: Wu, Lijun
; TITLE OF INVENTION: No. 6319675el Antibodies and Ligands for "Bonzo"
; FILE OF INVENTION: ChemoKine Receptor
; FILE REFERENCE: 1855.1070-000
; CURRENT APPLICATION NUMBER: US/09/449,437A
; CURRENT FILING DATE: 2001-01-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-449-437A-1

Query Match 54.3%; Score 679.2; DB 4; Length 1029;
Best Local Similarity 81.2%; Pred. No. 4.1e-215;
Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

Qy 166 AGGAGAACAAACGCTTCTTAAAGTTCAAGGAGGTCTTTTGGCCCTGTGTACTCTGGTAG 225
Db 62 AGGAGCATCAAGACTTCTCGAGTTCAAGAGGTCTTTTGGCCCTGATGATATATATATACACA 121
Qy 226 TGTTGTCTTTGGACTGCTAGGAACCTCCTGTTCTGATATATATATATATATATATATACACA 285
Db 122 TGTTGTCTGTTGGTCTGGGGAACCTCTCTGTTGCTGGTGCATATCATCTTCTACCATTA 181
Qy 286 AGCTGAGGACTCTGACAGATGTGTTTCTGCTGAACCTTGCCCTGGCTGACCTGGTGTG 345
Db 182 AGTTGCAGAGCTTGACCGATGTGTTCTGTTGTAACCTTACCCTGGCTGACCTGGTGTG 241
Qy 346 TCTGTACTCTGCCCTTTTGGGCTATGACGACACCTATGATGGGTCTTTGGCACAGTCA 405
Db 242 TCTGCACTCTGCCCTTCTGGGCTATGACGACACCTATGATGGGTCTTTGGCCAGGTCA 301
```

```
Qy 406 TGTGCAAAACTCTTCGAGGATGTATACAAATGAACCTTCTAGCTGTCCATGCTACTCTCA 465
Db 302 TGTGCAAAAGAGCCTACTGGGCATCTACACTATTAACTTTACACGTCCATGCTATCCTCA 361
Qy 466 CTTGATCACAAGTGGATCGTTTTCATTGTAGTGGTCCAGGCTACCAAGGCCCTTCAACCGGC 525
Db 362 CTTGATCACAAGTGGATCGTTTTCATTGTAGTGGTTCAGGCCACCAAGGCCCTACCAACGC 421
Qy 526 AGCTAAAGTGAAGATCTGGGGCCAAAGTCATTTGCTGCTCATTTGGTGGTGTCTCCCTGT 585
Db 422 AAGCAAGAGGATGACCTGGGGCAAGTTCACAGCTTGTCTCATCTGGGTGATATCCCTGC 481
Qy 586 TGGTTTCTTTCGCCACAGATCATCTATGSCCATGTTCAAGATATTGACAAGCTTATCTGTC 645
Db 482 TGGTTTCTTTCGCCCAAAATTAATCTATGSCAAATGCTTTAATCTCGACAGCTCATATGTG 541
Qy 646 ASTACCAAGTGAAGATATCCACTATGTTCTTGTATATACAGATGACTTGGGGTCT 705
Db 542 GTTACCATGACGAGGCAATTTCCACTGTGTGTTCTTGCACCCAGATGACACTGGGTCT 601
Qy 706 TCTTGCATTTGCTCACTATGATTTCTGTCTACTGAGCATTATCAAGACCTTGTCTCATG 765
Db 602 TCTTGCCTACTGCTCACCATGATTTGTCTGCTATTGATCATTAATCAAAACACTGCTTTCATG 661
Qy 766 CTCGAAACTTCCAGAAACACAAATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTC 825
Db 662 CTGGAGGCTTCCAGAAACACAGATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTC 721
Qy 826 TGCTGACCCAGACACCTTCAACCTTGGCCATGTTAAATCCAAAGTACAGCTGGGAGTACT 885
Db 722 TGCTGACCCAGATGCCCTTCAACCTCATGAAGTTTCATCCGAGCACACACTGGGAATACT 781
Qy 886 ATACCATTAACAGCTTTAAGTATGCTCATGCTAGTGACAGGCTATAGCATCTTTCGGG 945
Db 782 ATGCCATGACGAGCTTTCATCTACCATCATGTTGACAGAGGCCATCGCATCTGAGG 841
Qy 946 CTTGCTTAAACCTGTACTTTTATGCTTTTGTGCTTAAAGTTCCGGAAAGACCTCTGGA 1005
Db 842 CTTGCTTAAACCTGTGCTCTATGCTTTGTCAGCTGAAGTTTCGAAAGAACTTCTGGA 901
Qy 1006 AACTATCAAGGATATGCTGCTCTCTACCTGGGAGTCTCAAGTCAATGGAAGTCTT 1065
Db 902 AACTTGTGAAGACATGTTGTGCTTCCCTTACCTTGGGCTCTCATCAATGGAATCTT 961
Qy 1066 CTGAGGACAGTTCGAAGACTTGTCTGCTTCCCAATGTAGAGACCAAGTATGTTCC 1125
Db 962 CTGAGGACAAATTCGAAGACTTTTCTGCTTCCCAATGTAGAGACCAAGTATGTTCC 1021
Qy 1126 AATTGTAG 1133
Db 1022 AGTTATAG 1029

RESULT 5
US-09-275-384B-4
; Sequence 4, Application US/09275384B
; Patent No. 6232084
; GENERAL INFORMATION:
; APPLICANT: MACPHER, COLIN HOUSTON
; APPLICANT: MOORES, KITTY
; TITLE OF INVENTION: NEW USE
; FILE REFERENCE: GH-31106
; CURRENT APPLICATION NUMBER: US/09/275,384B
; CURRENT FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 980677.2
; PRIOR FILING DATE: 1998-03-27
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
```

US-09-275-384B-4

```
Query Match      54.1%; Score 676.8; DB 3; Length 1026;
Best Local Similarity 81.2%; Pred. No. 2.5e-214;
Matches 783; Conservative 0; Mismatches 181; Indels 0; Gaps 0;

QY 166 AGAGAACAAACGCTTCCTAAAGTTCAGAGAGTCTTTTGGCCCTGTGTACTCTGGTAG 225
DB 62 AGAGCATCAAGACTTCTCGAGTTCAGCAAGTCTTTTGGCCCTGTGTACTCTGGTAG 121
QY 226 TGTGTGTCTTTGGACTGTCTAGAACTCCCTGGTCTGATATATATACATTTCTACACAGA 285
DB 122 TGTGTGTCTGTGTCTGTGGGAACTCTGTGTGTCTGTGTATATCAATCTTCTACACATA 181
QY 286 AGCTGAGGACTCTGACAGATGTGTGTCTGAACTTCCCTGTGTGTGTGTGTGTGTGT 345
DB 182 AGTTGAGAGGCTGACGAGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 241
QY 346 TCTGTACTCTGCCCTTTTGGGCTATGAGGACCTATGAGTGGTCTTTGGCAGATCA 405
DB 242 TCTGCACCTCTGCCCTTTTGGGCTATGAGGACCTATGAGTGGTCTTTGGCAGGTC 301
QY 406 TGTGCAAACTCTTGGAGCATGTATACAACTTCTTACGTGTCCATGTCTCACTCTCA 465
DB 302 TGTGCAAGAGCTTACTGGGCACTTACATTTAACTTCTACAGTCCATGTCTCACTCTCA 361
QY 466 CTGTGATCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 525
DB 362 CTGTGATCAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 421
QY 526 AGCTTAAGTGAAGATCTGGGGCCAACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 585
DB 422 AGCCCAAGAGGATGACCTGGGGCAAGTCCAGAGTGTGTGTGTGTGTGTGTGTGTGT 481
QY 586 TGTGTCTTTGGCCAGAGTATCTATGAGGCAATGTTCAGATATTCAGAACTTATCTGTCT 645
DB 482 TGTGTCTTTGGCCCAAAATATCTATGAGGCAATGTTCAGATATTCAGAACTTATCTGTCT 541
QY 646 AGTACCAAGTGAAGATATCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 705
DB 542 GTTACCATGACAGGCAATTTCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 601
QY 706 TCTGCCATTTGCTACTATGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 765
DB 602 TCTGCCATTTGCTACTATGATTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 661
QY 766 CTCGAACTTCCAGAGCACAAATCTTAAGATCATCTTCTGTGTGTGTGTGTGTGTGTGT 825
DB 662 CTGGAGGCTTCCAGAGCACAGATCTTAAAGATCATCTTCTGTGTGTGTGTGTGTGTGT 721
QY 826 TGTGACCCAGACACCTTCAACCTTGGCATGTATTAATCAAGTACAAAGTGGAGTACT 885
DB 722 TGTGACCCAGATGCTTCAACCTTCAAGTTCATTCGAGACACACACTGGGATACT 781
QY 886 ATACCATCAACGCTTTAAGTATGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 945
DB 782 ATGCCATGACGAGTTCCTACATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 841
QY 946 CTGTGCTTAACTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1005
DB 842 CTGTGCTTAACTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 901
QY 1006 AACTTATGAAGATATGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1065
DB 902 AACTTGTGAAGATATGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 961
QY 1066 CTGAGGACAGTTCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1125
DB 962 CTGAGGACAAATTCGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1021
QY 1126 AATT 1129
DB 1022 AGTT 1025
```

```
RESULT 6
US-09-116-498-1
; Sequence 1, Application US/09116498
; Patent No. 6251582
; GENERAL INFORMATION:
; APPLICANT: Littman, Dan R.
;            Deng, Hongkui
;            Unutmaz, Derya
;            Ramani, Vireet N.K.
; TITLE OF INVENTION: NOVEL ALTERNATIVE G-COUPLED RECEPTORS
;                     ASSOCIATED WITH RETROVIRAL ENTRY INTO CELLS, METHODS OF
;                     IDENTIFYING THE SAME, AND DIAGNOSTIC AND THERAPEUTIC USES
;                     THEREOF
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
;            Floor
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/116,498
; FILING DATE: 16-Jul-1998
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-009
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1026 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-116-498-1

Query Match      51.8%; Score 648; DB 3; Length 1026;
Best Local Similarity 80.8%; Pred. No. 9.4e-205;
Matches 805; Conservative 0; Mismatches 185; Indels 6; Gaps 4;

QY 141 TTCAACAATTCACAGTGTATACAGCC--AGGAGAACAAACGCTTCTTAAAGTTCAAGGAG 197
DB 34 TTACAGAGTTTCAATGACACAGCAGGAGGAGCATCAAGACTTCTGCAAGTTCAAGAG 93
QY 198 GTCTTTTGGCCCTGTGTGTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 257
DB 94 GTCTTTTGGCCCTGTGTGTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 153
QY 258 GTTCTGATATATACATTTTCTACAGAGCTGAGGACTCTGACAGATGTGTGTGTGTGTGT 317
DB 154 G-TGTGTGTATATCCATCTTCTACCAATAGTTTCAGAGCCTGACGATGTGTGTGTGTGT 211
QY 318 AACTTGGCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 377
DB 212 AACCTACCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 271
```

QY 378 ACCTATGAGTGGGCTTTGGCACAGTCATGCGAAAACTCTTCAGGCAATGTATACAATG 437
 Db 272 ATCCATGAATGGGTGTTGGCCAGGTTCATGTGCAAGAGCCCTAGG-GGCATCTACACTATT 330
 QY 438 AACTTCTACGTGTCATGCTCACTCTCACTGATCAGCTGATCAGCTGGATCGTTTCATTGTAGTG 497
 Db 331 AACTTCTACAGTCCATGCTCATCTCCTCCTGATCAGCTGATGATGCTGATGCTGATGCTGATG 390
 QY 498 GTCCAGGCTACCAAGGCTTCAACCGGCAAGCTTAAGTGAAGATCTTGGGCCCAAGTCAAT 557
 Db 391 GTTAAGGCCCAAGGCTTCAACCAAGCAAGCCAAAGAGATGACCTTGGGCAAGGTCAAC 450
 QY 558 TGCTTGCTCAATTTGGGTGGTCTCCCTGTTGGTTCCTTTCGACAGATCATCTATGAGCCAT 617
 Db 451 AGCTTGCTCATCTGGGTGATATCCCTGCTGTTCTTTCGCCCCAAATTAATCTATGCAAT 510
 QY 618 GTTCAAGATATTGAAGCTTATCTGTCAGTACACAGTACAGGAGATATCCATATGGT 677
 Db 511 GTCTTTAATCTCGAAGCTCATATGTTGTTACCATGACGAGGCAATTTCCACTGTGGTT 570
 QY 678 CTGTTATACAGATGACTCTGGGGTCTCTCTGCGCATGCTCATATGATTTCTGTGCTAC 737
 Db 571 CTTCGCCACCCAGATGACACTGGGGTCTCTCTGCGCATGCTCATATGATTTCTGTCTAT 630
 QY 738 TCAGGCAATATCAAGACCTTCTTCATGCTCGAAACTTCCAGAAAGCACAATCTCTAAAG 797
 Db 631 TCAGTCATAATCAAAACACTGCTTCATGCTGGAGGCTTCCAGAAAGCACAATCTCTAAAG 690
 QY 798 ATCATCTTCTGTTAGTGGTCTGCTGCTGACCCAGACACCTTCAACCTTGCATG 857
 Db 691 ATCATCTTCTGTTAGTGGTCTGCTGCTGACCCAGATGCTTCAACCTTGCATG 750
 QY 858 TTAATCCAAAGTACAAGCTGGAGTACTATACCATACCAAGCTTAAAGTATGCCATCGTA 917
 Db 751 TTATCCCGCAGCACACTGGGAATATATGCCATGACCAAGCTTCACTACACCATCATG 810
 QY 918 GTGACAGAGGTATAGCATATTTGGGGCTTGCCTTAACCTGCTGCTATTTATGCTTTGTT 977
 Db 811 GTGACAGAGGCTATGCTATCTGAGGCTGCTGCTTAACTGCTGCTATGCTTTGCT 870
 QY 978 GGCTTAAAGTCCGGAAGAGCTCTGGAACATTTATGAAGATATCGGCTGCTCTCTCAC 1037
 Db 871 AGCCTGAAGTTTCGAAAGAACTTCTGAAAGCTTGTGAAGAGACATTTGGTTCCTTCTAC 930
 QY 1038 CTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTTCAAGACTTCTTCTGCTCTC 1097
 Db 931 CTTCGGGTCTCAATCAATGGAATCTTCTGAGGACAAATTCAGAGACTTTTCTGCTCTC 990
 QY 1098 CACAATGTAGACACACCATGATTTCCAAATGTAG 1133
 Db 991 CACAATGTGAGGCGCACAGCATGTTCCAGTTATAG 1026

RESULT 7
 US-09-170-496D-23
 ; Sequence 23, Application US/09170496D
 ; Patent No. 6555339
 ; GENERAL INFORMATION:
 ; APPLICANT: Behan, Dominic P.
 ; APPLICANT: Chalmers, Derek T.
 ; APPLICANT: Liaw, Chen W.
 ; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
 ; FILE OF INVENTION: Receptors
 ; FILE REFERENCE: AREN-0040
 ; CURRENT APPLICATION NUMBER: US/09/170,496D
 ; CURRENT FILING DATE: 1998-10-13
 ; NUMBER OF SEQ ID NOS: 294
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 23
 ; LENGTH: 1074
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-170-496D-23

Query Match 14.7%; Score 183.4; DB 4; Length 1074;
 Best Local Similarity 52.9%; Pred. No. 2.3e-50;
 Matches 490; Conservative 0; Mismatches 401; Indels 36; Gaps 3;
 QY 146 CAATTCACAGTATCAACAGCCAGGAGAACAAACGGCTTCTTAAAGTTCAAGGAGGTCTTTT 205
 Db 57 CAATTCACAGTATCAACAGCCAGGAGAACAAACGGCTTCTTAAAGTTCAAGGAGGTCTTTT 116
 QY 206 GCCCTGCTGTACCTGGTGTGTTTGTCTTTGGACTGCTAGGAAACTCCCTGGTCTGTAT 265
 Db 117 CCCACCTTGTACTGGCTCGTGTTCATCGTGGGTGCTTGGGCAACAGTCTTGTATTCTCT 176
 QY 266 TATATACATTTTACCAAGAGCTGAGGACTCTGACAGATGTGTTTCTGCTGACTGCTGCC 325
 Db 177 TGTCTACTGTGACTGCAAGAGTGAAGACCATGACCGACATGTTCTTTTGAATTTGGC 236
 QY 326 CTTGCTGACCTGGTGTGTTGTCTGTACTCTGCCCTTTTGGGCTTATGACGACCACTATGA 385
 Db 237 AATGCTGACCTCTCTTTCTTGTCTCTCTTCTTGGGCCATTTGCTGCTGTGACCA 296
 QY 386 GTGGTCTTTGGCACAGTCAATGTGCAAAACTCTTCGAGGCAATGTATACAATGACTTCTA 445
 Db 297 GTGGAAGTTCAGACCTTTCATGTGCAAGGTGCTCAACAGCAATGTACAGATGAACACTCTA 356
 QY 446 CGTGTCCATGCTCACTCTCACTGATGATGCTGTTTCTTGTAGTGGTCTCAGGC 505
 Db 357 GAGCTGTGTGTTGCTGATCATGTGATCAGCGTGGACAGGTACATTTGCCATTTGCCAGGC 416
 QY 506 TACCAGGCTTTCAACCGGAGGCTTAAAGTGAAGATCTGGGGCCCAAGTCAATTTGCTTGT 565
 Db 417 CATGAGACACATATCTTGGAGGAGAAAGGCTTTTGTACAGCAAAATGGTTCCTTTTAC 476
 QY 566 CATTTGGGTGCTCTCCCTGTTGTTTCTTTCGCCACAGATCATCTATGCCATGTTTCAAGA 625
 Db 477 CATCTGGGTATTTGGCAGCTGCTCTCTGATCCCAAGATCTTTATACAGCAAAATCAAGGA 536
 QY 626 T-----ATTGCAAGCTTATCTGTCACTCAACAGTGGAGATATCCACTATG-- 674
 Db 537 GGAATCCGGCATTTGCTATCTGCACCATGTTTACCTTAGCGATGAGAGCACCAAACTGAA 596
 QY 675 -----GTTCTGTTATACAGATGACTCTGGGGTCTTCTCTGCCATTTGCTCACTATGAT 727
 Db 597 GTCACTGCTGTGACCTTGAAGGTCAATTTGGGGTCTTCTCTTCCCTTCGTTGCTATGGC 656
 QY 728 TCTGTGCTACTCAGGCATTTCAAGACCTTGTCTTCACTGCTCGAAACTTCCAGAGCAAA 787
 Db 657 TTGCTGCTATACCATCATTCATCACCCCTGATACAGCCCAAGAGTCTTCCAAAGCAAA 716
 QY 788 ATCTCTAAGATCATCTTCTTGTAGTGGCTGTGTTCTGCTGACCCAGACACCTTCAA 847
 Db 717 AGCCCTAAAAGTGACCATCACTGTCTGACCGTCTTGTCTGTCTCAGTTTCCCTCAAA 776
 QY 848 CTTTCCCATGTTTAAATCCAAAGTACAAAGCTGGGAGTACTATATACCATACCAAGCTTAAAGTA 907
 Db 777 CTGCAATTTTGTGTGAGACCATTTGACGCTATGCCATGTTTCTTCCACTGTGCGGT 836
 QY 908 TGCCA-----TCGTAGTGCAGAGGCTATAGCATACTTTCGGGCTTG 949
 Db 837 TTCCACCAACATTTGACATCTGCTTCCAGGTCAACCCAGACCATCGCTTCTTCCACAGTTG 896
 QY 950 CTTTAACTCTGTACTTATGCTTTTGTGGCTTAAAGTTCGGGAAGAACGCTTCGAAACT 1009
 Db 897 CTTGAACCTGTTCTCTATGTTTTTGTGGTGAGAGATCCCGCGGATCTCTGTAAGAAC 956
 QY 1010 TATGAAGGATATCGGCTCCTCTCTCA 1036
 Db 957 CCTGAAGAACTGGGTTGATCAGCA 983

RESULT 8
 US-09-266-464-1
 ; Sequence 1, Application US/09266464

QY 531 AAGTGAAGATCTGGGGCCCAAGTCATTTGCTTGTCTCATTTGGGTGGTCTCCCTGTGGTT 590
Db 1139 CCGTGTCTCTCATCAGCAAGTGTCTGTGTGGGCACTCTGGATCTGGCCCTCTTCCTC 1198
QY 591 TCTTTGCCACAGATCATCTATGGCCATGTTCAAGATA-----TTGACA 633
Db 1199 TCATCCCGGAGCTCTCTACAGCGGCTCCAGAGACAGCGGGGAGGACAGCTGAGA 1258
QY 634 AGCTTATCTGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 693
Db 1259 TGCTCACTGGTCACTGCCCCAAGT--GGAGGCGCTTGATCACCATCCAAAGTGGCCCCAGATGG 1316
QY 694 CTCTGGGTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 753
Db 1317 TTTTGGGTTCTAGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1376
QY 754 CTTGCTTCTATCTCGAACTTCCAGAGACCAAAATCTCTAAAGATCATCTTCTCTTGTAG 813
Db 1377 CTTGCTGCTGAGGACGCAACTTTGAGCGGAACAAGGCCATCAAGGTGATCATTCGCGTGG 1436
QY 814 TGGCTGTGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
Db 1437 TGGTAGTCTTCTAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1496
QY 853 CCATGTTAATCAAAAGTACAAGCTGGAGTACTATACCATCAACAGCTTTAAGTATGCCA 912
Db 1497 CCAACTTCAACATCACCATAGCAGCTGCTGCGAACCAGCAGCAGCTCAACATGGCT 1556
QY 913 TCGTAGTGACAGAGCTATAGATCATCTTTGGGGCTTGGCTTAACTCTGCTTATTTATGGCT 972
Db 1557 ATGAGCTCACCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1616
QY 973 TTGTTGGCTTAAAGTTCGGAAGAGCTGCTGGAACCTTATGAAGGATATCGGCTGCTCT 1032
Db 1617 TCATCGGCTGAAGTTCGACGAGCCTCTTCAAGCTCTTCAAGGACTTTGGGCTGCCCTCA 1676
QY 1033 CTC 1036
Db 1677 GCA 1680

RESULT 11

US-09-299-843A-23
; Sequence 23, Application US/09299843A
; Patent No. 6107475
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: No. 6107475e1 Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; City: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,843A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/088,337
; FILING DATE: 01-JUN-1998
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/153,848
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jill E. Uhl
; REGISTRATION NUMBER: 43,213
; REFERENCE/DOCKET NUMBER: 27866/32059B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2751 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1..691
; FEATURE:
; NAME/KEY: exon
; LOCATION: 692..1771
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 692..1768
; FEATURE:
; NAME/KEY: polyA signal
; LOCATION: 2341..2348
; US-09-299-843A-23

Query Match 14.5%; Score 181.2; DB 3; Length 2751;
Best Local Similarity 53.2%; Pred. No. 2.3e-49;
Matches 481; Conservative 0; Mismatches 383; Indels 40; Gaps 3;

QY 171 AACAAACGCTTCCTAAAGTTCAAGGAGTCTTTTGGCCCTGTGTACCTGGTGTGTTT 230
Db 779 AAGAGGATGTGGGAACITTAAGGCTGTCTGCTCTCATGTATTCTGTCATGTC 838
QY 231 GTCTTTGGACTGTAGAAACTCCCTGGTTCTGATATATATATATATATATATATATAT 290
Db 839 TTCTGGGCTGTCTCGCAACGGGCTGGTATACCTGACGTATATATATATATATATAT 898
QY 291 AGGACTCTGACAGATGTGTTTCTGTGCACTTCCCTGCTGCTGCTGCTGCTGCTGCTGCT 350
Db 899 AAGACATGACGATACCTACCTGCTCACTGCGGCTGGCAGACATCTCTTTTCTCTTA 958
QY 351 ACTCTGCTCTTTTGGGCTATGCAAGCACTATGAGTGGGTCTTTGGCAGATCATGTGC 410
Db 959 ATTCTTCTCTTCTGGGCTACAGCAAGCAAGTCTCTGGATCTTTGGGCTCTACCTGTGT 1018
QY 411 AAAACTCTTCGAGGATGTATACAACTTCTAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCT 470
Db 1019 AAGGCACTTTTGGCATATATAAGTCTTTTCAAGGGGATGCTGTGCTCTCTATGTC 1078
QY 471 ATCAGATGTGATGCTTTTCAATTGTAGTGTCCAGGCTACCAAGGCTTCAACCGGCAAGCT 530
Db 1079 ATCAGCATTGACCGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAG 1138
QY 531 AAGTGAAGATCTGGGGCCCAAGTCACTTGTCTCATTTGGGTGGTCTCCCTGTGGTT 590
Db 1139 CGCTGCTTCTCTATCAGCAAGCTGTCTGTGTGGGATCTGGATCTGGCCCTCTTCCTC 1198
QY 591 TCTTTGCCACAGATCATCTATGGCCATGTTCAAGATA-----TTGACA 633
Db 1199 TCATCCCGGAGCTGCTCTACAGCGGCTCCAGAGACAGCGGGGAGGACAGCTGAGA 1258
QY 634 AGCTTATCTGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 693
Db 1259 TGCTCACTGGTCACTGCCCCAAGT--GGAGGCGCTTGATCACCATCCAAAGTGGCCCCAGATGG 1316

Db	1497	CGAACTTCACATCACCAATAGCAGCTGCTGCGAACCGAAGCAGCTCAACATTGCCT	1556
Qy	913	TCGTAGTGACAGAGGCTATAGCATACTTTCGGGCTTCGCTTAACCCCTGTACTTTATGCCT	972
Db	1557	ATGAGCTCACCTACAGCGTGGCCTCCGTCCGCTGCTGGGTCAACCCCTTCTTGATGCCT	1616
Qy	973	TTGTTGGCTTAAAGTTCCGAAGAACCTCTGGAACCTTATCAAGGATATCGGCTGCCTCT	1032
Db	1617	TCATCGGCGTCAAGTTCCGAGCGACCTCTTCAAGCTCTTCAAGGACTTGGGCTGCCTCA	1676
Qy	1033	CTCA	1036
Db	1677	GCCA	1680

RESULT 13

```

PC/US93-11153-23
; Sequence 23, Application PC/TUS9311153
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: Novel Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 64
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Hicknell
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/11153

```

Query Match	14.5%	Score 181.2	DB 5	Length 2751
Best Local Similarity	53.2%	Pred. No. 2.3e-49		
Matches 481	Conservative 0	Mismatches 383	Indels 40	Gaps 3
Qy	171	AACAAACGCTTCCTCAAGATTCAAGAGAGGTCCTTTTGGCCCTGTGTACCTGGTAGTGTGTTT	230	
Db	779	AAGAAGGATGTGCGGAACCTTTAAGGCGCTGGTCTCTGCCTCATGTATTCTGTGCATCTGC	838	
Qy	231	GTCTTTGGACTGTAGGAACCTCCCTGGTCTCATTTATATACATTTTCTACAGAGACTG	290	
Db	839	TTCGTGGGCTGTCTGGCAACGGGCTGGTGATACTGACGTACATCTATTTTCAAGAGGCTC	898	
Qy	291	AGGACTCTGACAGATGTGTTTCTGTGCAACTTCCCTCGCTGACCTGGTGTGTTGTCTGT	350	
Db	899	AAGACCATGAGGATACCTACCTGCTCAACTGGCGGTGGACACATCCCTTTTCCCTCCTA	958	
Qy	351	ACTCTGCCCTTTTGGGCGCTATGACGGACCTATGAGTGGGTCTTTGGCAGAGTCATGTGC	410	
Db	959	ATTCTTCCCTTCTGGGCGCTACACGGAAGCAAGTCTGGATCTTTGGCGTCTACCTGTGT	1018	
Qy	411	AAAACTCTCGAGGCGATGTATACAAATGAACCTTCACTGTCTCCATGTCTCACTTCACCTGC	470	
Db	1019	AAGGGCATCTTTGGCATCTATAAGTTAAGCTTCTTCAAGCGGATGCTGTGCTCTCATGC	1078	
Qy	471	ATCACAGTGAATCGTTTCATTGTAGTGGTCCAGGCTACCAAGGCTTTCACCGGAGGCT	530	
Db	1079	ATCAGCAATTACCGGCTACGTAAGCCATCGTCCAGGCGCGTGTGCGGTCAATGCCACCGGCC	1138	
Qy	531	AAGTGAAGATCTGGGGCCAAAGTCATTGTCTTCTCATTTGGGTGGTCTCCCTGTGTGTT	590	
Db	1139	CGCGTCTTCTCATCAGCAAGCTGTCTCTGTGTGGCATCTGGATGTGGCGCTCTTCTCCTC	1198	
Qy	591	TCCTTTGCCACAGATCATCTATATGGCCATGTTCAAGATA-----TTGACA	633	
Db	1199	TCCATCCGGAGCTGTCTACAGCGGCTCCAGAAGACAGCGCGAGGACACGCTTGAGA	1258	
Qy	634	AGCTTATCTGTCAGTACCAAGTGGAGATATCCACTATGGTCTTGTGTTATACAGATGA	693	
Db	1259	TGCTCACTGTCTAGTGCCCAAGT--GGAGGCTTGTATCACCATCCAAGTGGCCAGATGG	1316	
Qy	694	CTCTGGGGTTCTTCTTGCCATTTGCTCACTATGATTTCTGTGCTACTCAGGCAATTTACAGA	753	
Db	1317	TTTTTGGGTTCTTAGTGGCTATGCTGGCTATGAGTTTCTGTACCTCATTTATCATCCGTA	1376	
Qy	754	CTTGTCTCATGCTCGAAACTTCCAGAAGCACAAATCTCTAAAGATCATCTTCTTGTAG	813	
Db	1377	CTTGTCTCAGGACGCAACTTTAGCGGNAACAAGGCCATCAAGGTGATCATTTGCCGTGG	1436	
Qy	814	TGGGTGTTCTCTGTGACCCAGACACCTTTCAACC-----TTG	852	
Db	1437	TGTTAGTCTTTCATAGTCTTCCAGCTGCCCTACAAATGGGTGGTCTGGCTCAGACGGTGG	1496	
Qy	853	CCATGTTATCCAAAGTACAAAGTGGGAGTACTATACCAATACCAAGCTTTTAAGTATGCCA	912	
Db	1497	CCAACTTCAACTACCAATAGCAGTGTGCGAACCAGCAGCAGCTCAACATTGCCT	1556	
Qy	913	TCGTAGTGACAGAGGTATAGCATATCTTTCGGGCTTGCCTTAAACCTGTACTTTATGCTCT	972	
Db	1557	ATGACGTCACCTACAGCTGGGCTCGGCTCGCTGCTGCGTCAACCCCTTTCTTGTATGCTCT	1616	
Qy	973	TTGTTCGCTTAAAGTTCCGGAAGAACGTCGGAACCTTATGAAGGATATCGGCTGCTCT	1032	
Db	1617	TCATCGGCTCAAGTTCCGACGGACCTCTTCAAGTCTTCAAGGACTTGGGCTGCTCA	1676	
Qy	1033	CTCA	1036	
Db	1677	GCCA	1680	

RESULT 14
US-09-170-496D-175
; Sequence 175, Application US/09170496D
; Patent No. 6555339

; GENERAL INFORMATION:

; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 175
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-170-496D-175

Query Match 14.3%; Score 178.6; DB 4; Length 1074;
Best Local Similarity 52.5%; Pred. No. 8.9e-49;
Matches 487; Conservative 0; Mismatches 404; Indels 36; Gaps 3;

Qy 146 CAATCCAGTGATACAGCCAGGAGACAAACCTTCCTAAAGCTTCAAGAGGTCTTTT 205
Db |||||
Qy 57 CAATCCAGTGATACAGCCAGGAGACAAACCTTCCTAAAGCTTCAAGAGGTCTTTT 205
Db |||||
Qy 206 GCCCTGTGTACTGCTGTAGTGTGTTTCTTTGGACTGCTAGGAAACTCCCTGGTGTCTGAT 265
Db |||||
Qy 117 CCCACCTTGTACTGGCTGTGTTTCATCGTGGTGGCTTGGCAACAGTCTTGTATCCT 176
Db |||||
Qy 266 TATATACATTTTACACAGAGCTGAGGACTCTGACAGATGTTTCTGCTGACTGCTGCC 325
Db |||||
Qy 177 TGCTACTGTACTGCAAGAGTGAAGACCATGACCGACATGTCCTTTGAATTTGCG 236
Db |||||
Qy 326 CCTGCTGACTGTGTTTCTGTACTCTGCTTGTGGCCCTTTGGGCCCTATGCGAGCACTATGA 385
Db |||||
Qy 237 AATGCTGACTCCTCTTCTGCTACTCTTCCCTCTGGGCCATTTGCTGCTGACCA 296
Db |||||
Qy 386 GTGGGCTTTTGGCAAGTATGTCGAAACTCTTCGAGGATGATATACAAAGTAACTCTTA 445
Db |||||
Qy 297 GTGGAAGTTCAGACCTTCATGTGCAAGTGTGCAACAGCATGTACAAGATGAATCTTA 356
Db |||||
Qy 446 CGTGCTACTCTCACTCACTGCTATCAGTCAAGTGGATGTTTCAATGTAGTGTGCTCAGGC 505
Db |||||
Qy 357 CAGCTGTGTGTGATGATGTCATGTCATGTCAGTGGAGAGTATTTGCCATTTGCCAGGC 416
Db |||||
Qy 506 TACCAAGGCTTCAACCGGAGGCTAAGTGGAGAGTCTGGGGCAAGTCAATTTGCTTGT 565
Db |||||
Qy 417 CATGAGAGCACATATTGGAGGGAGAAAGGCTTTTGTACAGCAAAATGTTTGTCTTAC 476
Db |||||
Qy 566 CATTTGGTGGTCTCCCTGTGTTTCTTTGCCACAGATCATCTATGGCCATGTTCAAGA 625
Db |||||
Qy 477 CATCTGGTATTTGGCAGCTGCTCTGTGATCTCCAGAAATCTTATACAGCAAAATCAAGA 536
Db |||||
Qy 626 T-----ATTGACAAGCTTATCTGTCAATACCAAGTACCAAGTACCAAGTATCCACTATG-- 674
Db |||||
Qy 537 GGAATCGGATTTGCTATCTGCACATGTTTACCCTAGCGATGAGAGCAACAACTGAA 596
Db |||||
Qy 675 -----GTTCTGTTATACAGATGACTTGGGGTCTTCTGCTGCTATGCTACTATGAT 727
Db |||||
Qy 597 GTCAGCTGTCTTGACCTGAGGTCATCTTCTGGGGTCTTCTTCCCTTCTGCTGCTGCTG 656
Db |||||
Qy 728 TCTGTGCTACTCAGGCAATATCAAGACCTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 787
Db |||||
Qy 657 TTGCTGTATACCATCATCATCATCACCCCTGATACAGCAAGAGTCTTCCAGACAA 716
Db |||||
Qy 788 ATCTCTAAAGATCATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 847
Db |||||
Qy 717 AGCCAAGAAAGTGACCATCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 776
Db |||||
Qy 848 CCTTGGCCATGTATATCAAAAGTCAAGCTGGAGTACTATACCAATACCAAGCTTTAAGTA 907
Db |||||
Qy 777 CTGCAATTTTGTGTGTCAGACCATTTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 836
Db |||||

Qy 908 TGCCA-----TCGTAGTGACAGAGCTATAGCATACTTTGGGCTTG 949
Db |||||
Qy 837 TTCCACCAACATTGACATCTCTTCCAGGTCACCCAGACCATCGCCTTCTTCCACAGTTG 896
Db |||||
Qy 950 CCTTAACCCCTTACTTATGCTTTGTTGCTTAAAGTTCGGAAGAACGCTCTGGAACCT 1009
Db |||||
Qy 897 CTTGAACCCCTTCTCTATGTTTGTGGTGAAGATTCGCCGGGATCTCTGTGAAC 956
Db |||||
Qy 1010 TATGAAGATATCGCTGCCTCTCTCA 1036
Db |||||
Qy 957 CTTGAAGAACTTGGTTCATCAGCCA 983
Db |||||

RESULT 15

US-09-299-843A-65
; Sequence 65, Application US/09299843A
; Patent No. 6107475
; GENERAL INFORMATION:
; APPLICANT: Godiska, Ronald
; APPLICANT: Gray, Patrick W.
; APPLICANT: Schweikart, Vicki L.
; TITLE OF INVENTION: No. 6107475el Seven Transmembrane Receptors
; NUMBER OF SEQUENCES: 66
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &
; ADDRESSEE: Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/299,843A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/088,337
; FILING DATE: 01-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/153,848
; FILING DATE: 17-NOV-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/977,452
; FILING DATE: 17-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jill E. Uhl
; REGISTRATION NUMBER: 43,213
; REFERENCE/POCKET NUMBER: 27866/32059B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; TELEX:
; INFORMATION FOR SEQ ID NO: 65:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2085 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 177..1310
; US-09-299-843A-65

Query Match 14.2%; Score 177.8; DB 3; Length 2085;
Best Local Similarity 52.9%; Pred. No. 2.6e-48;
Matches 477; Conservative 0; Mismatches 387; Indels 37; Gaps 3;

QY 171 AACAAACGCTTCTTAAGTCTCAAGGAGGTCTTTTGGCCCTGTGTGTACCTGGTAGTGT 230
|||
Db 324 AAGAGGATGTGGGAACCTTTAAGCCCTGGTTCCTGCTCTCATGTATTCGTGTCATCTGC 383
QY 231 GTCTTTGACTGTAGGAACCTCCCTGCTTCTGATATATATATATTTCTACCAAGCTG 290
|||
Db 384 TTCGTGGGCTGCTCGGAACGGGCTGGTGTACTGACGTACATCTATTTCAAGAGGCTC 443
QY 291 AGGACTCTGACAGATGTCTTCTGTGAACCTTGCCTGGCTGACCTGGTGTGTGTGT 350
|||
Db 444 AAGACCATGACGGATACCTCTGCTCAACCTGGCCGTGGCAGACATCTTTTCTCCTA 503
QY 351 ACTCTGCTCTTTGGGCTATGACGACCTATAGTGGTCTTTGGCAGATCATGTGC 410
|||
Db 504 ATTCTTCCCTTCTGGGCTCTAGCGAAGCCAAAGTCTTGGATCTTTGGGCTTACCTGTGT 563
QY 411 AAAACTCTTCGAGGATGTATACATGAACCTTACGTGTCCATGCTCACTCTCACCTGC 470
|||
Db 564 AAGGCATCTTTGGCATCTATAGTTAAGCTTCTTACGGGATGCTGCTCTCTATGC 623
QY 471 ATCAGTGTGATCGTTTCATTTAGTGTGTCAGGCTACCAAGGCTTCAACCGGCAAGCT 530
|||
Db 624 ATCAGCATTTACCGCTAGCTAGCCATCGTCCAGGCCGTGTCGCGTCATCGGCACCGGCC 683
QY 531 AAGTGGAGATCTGGGCCAAGTCATTTGCTGTCTCATTTGGGTGGTCTCCCTCTTGGTT 590
|||
Db 684 CGCGTGTCTTCATCAGCAAGCTGTCCTGTGTGGGATCTGGATGCTGGCCCTCTTCTC 743
QY 591 TCTTTGCCACAGATCATCTATGGCCATGTTCAAGATA-----TTGACA 633
|||
Db 744 TCCATCCCGAGCTCTCTACAGCGCCCTCCAGAAGACAGCGCGGAGCACGCTGAGA 803
QY 634 AGCTTATCTGTACGATACACAGTGTAGGAGATATCCACTATGGTCTTGTGTATACAGATGA 693
|||
Db 804 TGCTCACTGGTCAGTGTCCCAAGT--GGAGGCTTTGATCACCATCCAAGTGGCCAGATGG 861
QY 694 CTCTGGGTTCTTCTGCTGCTTGTCTACTGATTTCTGCTACTCAGGCATTATCAAGA 753
|||
Db 862 TTTTGGGTTCTAGTGTCTATGCTGGCTATAGTTCGTCTACTCATTAATCATCCGTA 921
QY 754 CTTGCTTCACTGCTGAACCTTCCAGAAGCAAAATCTCTAAAGATCATCTTCTCTGTAG 813
|||
Db 922 CTTGCTCCAGGCACGCAACTTTGAGCGGAACAAGGCCATCAAGGTGATCATTTGCCGTGG 981
QY 814 TGGCTGTCTTCTGCTGACCCAGACCCCTTCAACCTTGCCTGCTTATCCAAAGTACAA 873
|||
Db 982 TGGTAGTCTTATAGTCTTCCAGCTGCCCTACAATGGGGTGGTCTGCTCAGACGGTGG 1041
QY 874 GCTGGGATGATATACCATTAACCACT-----TTAAGTATGCCATCG 915
|||
Db 1042 CCAACTTCAATACCATACCAATAGCAGCTCGGAACCCAGCAAGCAGCTCAACATTGCCATG 1101
QY 916 TAGTGACAGAGGCTATAGCATCTTTGGGGCTTGGCTTAACCCCTGACTTTATGCCCTTTG 975
|||
Db 1102 AGTCACTTACAGCCTCGCCTCCGTCGCTGCTGCTCAACCCCTTCTTTGTATGCCCTCA 1161
QY 976 TTGGCTTAAAGTTCCGGAAGACGCTCTGGAACCTTATGAAGATATCGGCTGCCTCTCTC 1035
|||
Db 1162 TCGGGCTCAAGTTCGCGAGGACCTCTTCAAGCTCTTCAAGGACTTGGGCTGCTCAGCC 1221
QY 1036 A 1036
Db 1222 A 1222

Search completed: September 29, 2004, 03:02:04
Job time : 125 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 29, 2004, 00:36:55 ; Search time 637 Seconds
(without alignments)
9947.020 Million cell updates/sec

Title: US-10-603-141-1

Perfect score: 1251

Sequence: 1 gaattcgacagaggcaga.....gtatcatggagagcgtgac 1251

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 337386 seqs, 2532474682 residues

Total number of hits satisfying chosen parameters: 6674772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US08A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707	56.5	2008	15	US-10-198-846-13566
2	705.4	56.4	1953	13	US-10-342-887-1299
3	705.4	56.4	1953	13	US-10-172-118-1299
4	705.4	56.4	1953	13	US-10-225-567A-347
5	704.8	56.3	1918	13	US-10-458-110-1
6	689.4	55.1	1029	9	US-09-852-156-5
7	685	54.8	1037	9	US-09-852-156-3
8	680.8	54.4	1029	15	US-10-083-168-3
9	679.2	54.3	1029	9	US-09-940-063-1
10	679.2	54.3	1029	15	US-10-174-293-1
11	679.2	54.3	1029	16	US-10-353-690-5
12	676	54.0	1029	9	US-10-083-168-64
13	648	51.8	1026	9	US-09-852-156-1
14	183.4	14.7	1074	15	US-10-251-385-23

15	183.4	14.7	2462	15	US-10-225-567A-240	Sequence 240, Appl
16	183.4	14.7	2577	9	US-09-903-377-1	Sequence 1, Appli
17	183.4	14.7	2577	9	US-09-952-385-1	Sequence 1, Appli
18	183.4	14.7	2577	10	US-09-966-755-1	Sequence 1, Appli
19	183.4	14.7	2577	14	US-10-000-739A-1	Sequence 1, Appli
20	183.4	14.7	2577	16	US-10-305-720-1095	Sequence 1095, Ap
21	180.2	14.4	1110	11	US-09-826-509-480	Sequence 480, App
22	178.6	14.3	1074	15	US-10-251-385-175	Sequence 175, App
23	177.8	14.2	1137	15	US-10-004-113-21	Sequence 21, Appl
24	177.8	14.2	2072	13	US-09-815-937-19	Sequence 19, Appl
25	177.8	14.2	2072	15	US-10-004-113-20	Sequence 20, Appl
26	169.6	13.6	30828	15	US-10-004-113-19	Sequence 19, Appl
27	163.6	13.1	1221	15	US-10-029-386-22664	Sequence 22664, A
28	163.6	13.1	2981	13	US-10-087-192-1319	Sequence 1919, Ap
29	163.6	13.1	36176	13	US-10-087-192-1318	Sequence 1918, Ap
30	162	12.9	1221	15	US-10-325-430-17	Sequence 17, Appl
31	162	12.9	1255	17	US-10-641-643-993	Sequence 993, App
32	162	12.9	1679	16	US-10-305-720-1097	Sequence 1097, Ap
33	162	12.9	2100	15	US-10-313-542-289	Sequence 289, App
34	162	12.9	3693	15	US-10-325-430-16	Sequence 16, Appl
35	161.6	12.9	1280	13	US-10-087-192-1916	Sequence 1916, Ap
36	161.6	12.9	1302	13	US-10-244-089-1	Sequence 1, Appli
37	161.6	12.9	1302	15	US-10-254-089-1	Sequence 1, Appli
38	161.6	12.9	34455	13	US-10-087-192-1915	Sequence 1915, Ap
39	161.4	12.9	1137	15	US-10-251-385-73	Sequence 73, Appl
40	161.4	12.9	1137	15	US-10-251-385-203	Sequence 203, App
41	161.4	12.9	1137	15	US-10-004-113-24	Sequence 24, Appl
42	161.4	12.9	1500	17	US-10-641-643-1465	Sequence 1465, Ap
43	161.4	12.9	2139	13	US-10-342-887-678	Sequence 678, App
44	161.4	12.9	2139	13	US-10-172-118-678	Sequence 678, App
45	161.4	12.9	2139	15	US-10-225-567A-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1

US-10-198-846-13566
; Sequence 13566, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steimann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13566
; LENGTH: 2008
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2008
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-13566

Query Match 56.5%; Score 707; DB 15; Length 2008;
Best Local Similarity 78.8%; Pred. No. 3.3e-225;
Matches 871; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

QY 141 TTCAACAATTCCAGTGTATACAGCC---AGGAGACAAACGGTTCTTAAGTTCAAGGAG 197
|||||
Db 144 TTTCAGCAGTTTCAATCATCAGCAGCAGGAGGAGCATCAAGACTTCTTCGAGTTCACGAG 203
|||||

198 GTCTTTTGGCCCTGTGTGTAACCTGGTAGTGTGTTGTTCTTTGGACTGCTAGGAACCTCCCTG 257
 204 GTCTTTTGGCCCTGTGTGTAACCTGGTAGTGTGTTGTTCTTTGGACTGCTAGGAACCTCCCTG 263
 258 GTTCTGATTTATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTTGTTCTGCTG 317
 264 GTGCTGCTATATCCATCTTCTACCAATAAGTTGACAGCCTGACGATGTTGTTCTGCTG 323
 318 AACTTGGCCCTGTGTGTAACCTGGTAGTGTGTTGTTCTTTGGACTGCTAGGAACCTCCCTG 377
 324 AACTTGGCCCTGTGTGTAACCTGGTAGTGTGTTGTTCTTTGGACTGCTAGGAACCTCCCTG 383
 378 ACCTATGAGTGGTCTTTGGACAGTCTATCTGCAAACTCTTGGAGCATGTATACAAGT 437
 384 ATCCATGAATGGGTGTTGGCCAGGTCTATGTGCAAGCCTACTGGGCATCTACATAT 443
 438 AACTTCTACGTGTCTACCTCTCACTCTCACCTGCATCACAGTGGATCGTTTCAATGTAGTG 497
 444 AACTTCTACAGTCTCATCTCTCACCTGCATCACAGTGGATCGTTTCAATGTAGTG 503
 498 GTCCAGCTACCAAGGCTTCAACCCGACGCTAAGTGGAGATCTGGGCGCAAGTCATT 557
 504 GTTAAGGCCCAAGGCTTCAACCCGACGCTAAGTGGAGATCTGGGCGCAAGTCATT 563
 558 TGCTTGTCTATTTGGTGGTCTCCCTGTTGGTCTTTTGGCCAGATCATCTATGCGCCAT 617
 564 AGCTTGTCTATTTGGTGGTCTCCCTGTTGGTCTTTTGGCCAGATCATCTATGCGCCAT 623
 618 GTTCAAGATTTGCAAGCTTATCTGTACATACACAGTGGAGATCTCCACTATGTTT 677
 624 GTCTTTAATCTCGAAGCTCATATGTGGTTTACCATGACGAGCAATTTCCACTGTGTT 683
 678 CTGCTTATACAGATCTCTGGGCTTCTCTGCAATGCTCACTATGTTCTGCTCTAC 737
 684 CTGCTTATACAGATCTCTGGGCTTCTCTGCAATGCTCACTATGTTCTGCTCTAC 743
 738 TCAGGCATTTCAAGACTGTTGCTTCACTGCTGCAATGCTCACTATGTTCTGCTCTAC 797
 744 TCAGTCATTTCAAACTGCTTCACTGCTGCAATGCTCACTATGTTCTGCTCTAC 803
 798 ATCATCTTCTGTTAGTGGTCTGTTCTGCTGCAATGCTCACTATGTTCTGCTCTAC 857
 804 ATCATCTTCTGTTAGTGGTCTGTTCTGCTGCAATGCTCACTATGTTCTGCTCTAC 863
 858 TTAATCCAAAGTACAAAGCTGGAGTACTATACCAATACCAAGCTTAAAGTATGCCATCGTA 917
 864 TTCAATCCGACGACACACTGGGAATCTATGCCATGACCAAGCTTCACTACACCATATG 923
 918 GTGACAGAGCTATAGATCTTTCGGCTGTTGCTTAACTGATCTTATGCTGCTTGT 977
 924 GTGACAGAGCTATAGATCTTTCGGCTGTTGCTTAACTGATCTTATGCTGCTTGT 983
 978 GGCCTTAAAGTTCGGAAGAACTGCTGAACTTATGAAGATATCGGCTGCTCTCTCAC 1037
 984 AGCTGAGTTTGAAGAACTTCTGGAATCTTGAAGATATGCTGCTGCTCTCTCAC 1043
 1038 CTGGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTTCCAAAGCTGTTCTGCTCTC 1097
 1044 CTGGGAGTCTCAATGGAAGTCTTCTGAGGACAGTTTCCAAAGCTGTTCTGCTCTCTC 1103
 1098 CACAATGTAGAGACCAAGTATGTTTCCAATTTGATGAGCCTTGGCACAATTTAGAGAAG 1157
 1104 CACAATGTAGAGACCAAGTATGTTTCCAATTTGATGAGCCTTGGCACAATTTAGAGAAG 1160
 1158 TTAATAACAGATTTCTAGGAGCATGCTGTATCATTTGGATGCAACAGAAAGCTTTGCT 1217
 1161 GCTGCTCTGGAATTTGCAAGGCTATGCTGCTCTCTGATGTTGGTGGTGGTGGTGGT 1220
 1218 TTATAGCATGTGGAGTATCATGGAGAG 1245
 1221 TTATAGCTTGGCATTTCTCATGGAGAG 1248

RESULT 2
 US-10-342-887-1299
 ; Sequence 1299, Application US/10342887
 ; Publication No. US20040058340A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Dai, Hongyue
 ; APPLICANT: He, Yudong
 ; APPLICANT: Linsley, Peter S.
 ; APPLICANT: Mao, Mao
 ; APPLICANT: Roberts, Christopher J.
 ; APPLICANT: Van 't Veer, Laura Johanna
 ; APPLICANT: Van de Vijver, Marc J.
 ; APPLICANT: Bernards, Rene
 ; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
 ; FILE REFERENCE: 9301-188-999
 ; CURRENT APPLICATION NUMBER: US/10342,887
 ; CURRENT FILING DATE: 2003-01-15
 ; PRIOR APPLICATION NUMBER: 60/298,918
 ; PRIOR FILING DATE: 2001-06-19
 ; PRIOR APPLICATION NUMBER: 60/380,710
 ; PRIOR FILING DATE: 2002-05-14
 ; PRIOR APPLICATION NUMBER: 10/172,118
 ; PRIOR FILING DATE: 2002-06-14
 ; NUMBER OF SEQ ID NOS: 2699
 ; SEQ ID NO 1299
 ; LENGTH: 1953
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-342-887-1299

Query Match 56.4%; Score 705.4; DB 13; Length 1953;
 Best Local Similarity 78.5%; Pred. No. 1.1e-224;
 Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;

QY . 141 TTCACAAATTCACGTATACAGCC---AGGAGAAACAAACGCTTCTTAAAGTTCAAGGAG 197
 Db 115 TTCACAGATTTCAATGACAGCAGCAGGAGGAGCATCAAGACTTCTGCGAGTTCAGCAAG 174
 QY 198 GTCTTTTGGCCCTGTGTGTAACCTGGTAGTGTGTTGTTCTTTGGACTGCTAGGAACCTCCCTG 257
 Db 175 GTCTTTTGGCCCTGTGTGTAACCTGGTAGTGTGTTGTTCTTTGGACTGCTAGGAACCTCCCTG 234
 QY 258 GTTCTGATTTATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTTGTTCTGCTG 317
 Db 235 GTGCTGTCATATCCATCTTCTACCAAGTTGACAGCCTGACGATGTTTCTGCTG 294
 QY 318 AACTTGGCCCTGGCTGACCTGCTGTTGTTCTGTACTCTGCTCTTGGGCTTATGCAAGC 377
 Db 295 AACCTACCCCTGGCTGACCTGCTGTTGTTCTGTGCTCTGCTCTGCTGCTATGCAAGC 354
 QY 378 ACCTATGAGTGGGTCTTTGGACAGATCATGTGCAAACTCTTGGAGCATGTATACATG 437
 Db 355 ATCCATGAATGGGTGTTTGGCCAGTCTATGTGCAAGAGCCTTCTGCGCATCTACATAT 414
 QY 438 AACTTCTACGTCCATGCTCACTCTCACTGTCATCACAGTGGATCGTTTCAATGTAGTG 497
 Db 415 AACTTCTACAGTCCATGCTCATCTCACTGTCATCACTGTCATCTGTCATGTTAGTG 474
 QY 498 GTCCAGGCTACCAAGGCTTCAACCCGACGCTAAGTGGAGATCTGGGCGCAAGTCATT 557
 Db 475 GTTAAGGCCCAAGGCTTCAACCCGACGCTAAGTGGAGATCTGGGCGCAAGTCATT 534
 QY 558 TGCTTGTCTATTTGGTGGTCTCCCTGTTGGTCTTTTGGCCAGATCATCTATGCGCAT 617
 Db 535 AGCTTGTCTATCTGGTGGTATCTCCCTGCTGCTGTTTCTTGGCCAAATATCTATGCAAT 594
 QY 618 GTTCAAGATTTGCAAGCTTATCTGTACGTACCAAGTGGAGAGATATCCACTATGCTT 677
 Db 595 GTCTTTAATCTCGAAGCTCATATGTGGTTTACCATGACGAGGCAATTTCCACTGTGTT 654
 QY 678 CTGTTATACAGATGACTCTGGGCTTCTTCTGCTGCAATGCTCACTATGTTCTGCTGCTAC 737
 Db 655 CTTCGCCCAAGATGACACTGGGTTCTTTTGGCCACTGCTTCAAGATGTTCTGCTAT 714

QY	738	TCAGGCATTATCAAGACCTTGCCTTCATGCTCGAAACTTCCAGAAGCACAATCTCTAAAG	797
Db	715	TCAGTCATAATCAAAAACACTGCTTCATGCTCGAGGCTTCCAGAAGCAGATCTCTAAAG	774
QY	798	ATCATCTTCCTGTAGTGGCTGTCTCTGCTGACCCAGACACCTTCAACCTTGCCTATG	857
Db	775	ATCATCTTCCTGGTGTAGTGGCTGTCTCTGCTGACCCAGATGCGCTTCAACCTCATGAAG	834
QY	858	TTAATCCAAAGTACAAGCTGGGAGTACTATACCATACACGCTTTAAAGTATGCCATGTA	911
Db	835	TTCATCCGCACACACACTGGGAATACTATGCCATGACCCAGCTTTCATACACCATCATG	894
QY	918	GTGACAGAGGCTATAGCAATCTTCGGGCTTGCCCTTAACCTGTACTTTATGCTTTGTT	977
Db	895	GTGACAGAGGCTATCGCATCTCAGGGCTTGCCTTAACCTGTGCTCTATGCTCTTGTG	954
QY	978	GGCTTAAAGTTCGGGAAGAACGCTCTGGAACACTTAATGAAGGATATCGGCTGCCTCTCAC	1037
Db	955	AGCCTGAAGTTTCGAAGAACTCTTGGAACCTGTGAAGGACATGGTGGCTCCCTTAC	1014
QY	1038	CTGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTCCTCAAGACTTGTTCGCTCC	1097
Db	1015	CTTGGGTCTCACATCAATGGAATCTTCTGAGGACAACTTCCAAAGACTTTTCTGCCTCC	1074
QY	1098	CACAATGTAGAGACCAACAGTATGTTTCCAATGTATAGGCTTGCACACACTTAGAGAAG	1157
Db	1075	CACAATGTGGAGGCCACAGCATGTTCCA---GTTATAGGCTTGCAGGGTTTCGAGAA	1131
QY	1158	TTAATAACAGAAATCTAGGACCATGGCTGTATCATTTGGATGCAACAAGAAAAGCTTTGC	1217
Db	1132	GCTGCTCTGGAAATTTGCAAGTCATGGCTGTGCTCTCTGTATGCTGGTGAAGCAGCTTGT	1191
QY	1218	TTATAGCAATGGAGTATCATGGAGAAG	1245
Db	1192	TTATAGCTTGGCATCTCTCATGGAGAAG	1219

RESULTS, T 3

```

RES001.3
US-10-172-118-1299
Sequence 1299, Application US/10172118
Publication No. US20030224374A1
GENERAL INFORMATION:
APPLICANT: Dai, Hongyue
APPLICANT: He, Yudong
APPLICANT: Linsley, Peter
APPLICANT: Mao, Mao
APPLICANT: Roberts, Chris
APPLICANT: Van 't Veer, Laura
APPLICANT: Van de Vijver, Marc
APPLICANT: Bernards, Rene
TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
FILE REFERENCE: 9301-175-999
CURRENT APPLICATION NUMBER: US/10/172,118
CURRENT FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 60/380,770
PRIOR FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2699
SEQ ID NO 1299
LENGTH: 1953
TYPE: DNA
ORGANISM: Homo sapiens
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: NM_006564
DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1299

```

Query Match 56.4%; Score 705.4; DB 13; Length 1953;
Best Local Similarity 78.5%; Pred. No. 1.1e-224;
Matches 870; Conservative 0; Mismatches 232; Indels 6; Gaps 2;
Qy 141 TTCACAAATTCACGTGATAACAGCC---AGGAGAACAAACGCTTCCTAAAGTTCAAGGAG 197

Db	115	TT	CAGCAGTTTTC	CAATGACAGCAGCAGCAGGAGGAGCATCAAGACTTCTCTGCAGTTTCAGCAAG	174
Qy	198	GT	CTTTT	TGGCCCTGTGTGTACCTGGTAGTGTTTGTCTTTGGACTGTCTAGAAACTCCCTG	257
Db	175	GT	CTTTCTGCCC	TGCATGTAACCTGGTGGTCTTTGTCTGTGCTGTGGTGGGAACTCTCTG	234
Qy	258	GT	TTCTGATATATACATTTTCTACAGAGACTGAGGACTCTGACAGATGTGTTCCTGCTG	317	
Db	235	GT	CTGTGTCATATCCATCTTCTACACCAAGTGTGCAGAGCCTGACGGAATGTGTTCCTGGTG	294	
Qy	318	AA	CTTGCCCTCGCTGACCTGGTGTGTGTCTGTACTCTGCCCTTTTGGGCGCTATGCAAGC	377	
Db	295	AA	CTACCCCTGGCTGACCTGGTGTGTGTCTGCACCTCTGCCCTCTGGGCGCTATGCAAGC	354	
Qy	378	AC	CTATGAGTGGGTCTTTGGCAGCAGTCAATGTGCAAAACCTCTTCGAGGCATGTATACAATG	437	
Db	355	AT	CCATGAATGGGTGTTTGGCCAGGTCATGTGAAGAGCCCTACTGGGCACTCTACACTATT	414	
Qy	438	AA	CTTCTACGTGTCCATGCTCACTCTCACCTGCATCACAGTGGATCGTTTCATTGTAGTG	497	
Db	415	AA	CTTCTACAGTCCATGCTCATCTCTCACCTGCATCACTGTGGATCGTTTCATTGTAGTG	474	
Qy	498	GT	CCAGGCTACCAAGGCCCTTCAACCGCAGGCTTAAGTGGAGATCTGGGCGCCAAAGTCATT	557	
Db	475	GT	TAAAGCCACCAAGGCCCTCAACACGAAAGCCAAAGAGGATGACCTGGGCGCAAGTCAAC	534	
Qy	558	TG	CTTCTCATTTGGGTGGTCTCCCTGTGTGGTTCTTTGGCCACAGATCATCTATGGCCAT	617	
Db	535	AG	TTGCTCATCTGGGTGATATCCCTGCTGGTTTCTTGGCCCAAAATATCTATGGCAAT	594	
Qy	618	GT	TCAAGATATTGACAAGCTTATCTGTCACTAGCACAGTGAAGGAGATATCCACTATGCTTT	677	
Db	595	GT	CTTTAACTCTCGACAAGCTCATATGTGGTTACCATGACGAGGCAATTTCCACTGTGTT	654	
Qy	678	CT	TGTTATACAGATGACTCTGGGGTCTTCTCGCATTGTCTACTATGATTTCTGTGCTAC	737	
Db	655	CT	TGCCACCCAGATGACCTGGGGTCTTCTTGGCACTGCTCCACATGATGTCTGTCTAT	714	
Qy	738	TC	AGGCAATTATCAAGACTTCTGCTTCATGCTCGAAACTTCCAGAAAGCAAAATCTCTAAAG	797	
Db	715	TC	AGTCATATCAAAACACTGCTTCATCTGGAGGCTTCCAGAAGCACAGATCTCTAAAG	774	
Qy	798	AT	CACTCTCTGTAGTGGCTGTCTCTGCTGACCCAGACACCCTTCAACCTTGCCTG	857	
Db	775	AT	CACTCTCTGTGTGGTGTGTCTCTGCTGACCCAGATGCCCTTCAACCTCATGAAG	834	
Qy	858	TT	ATCCAAAGTACAAGCTGCGAGTACTATACCAATACCAAGCTTAAAGTATGCCATGTA	917	
Db	835	TT	CATCCGAGCACACTGGGAATACTATGCCATGACACGCTTCTACTACCATCATG	894	
Qy	918	GT	CACAGAGCTATAGCATATTTTCGGGCTTGCTTTAAACCTGTACTTTTATGCTTTGTT	977	
Db	895	GT	CACAGAGCCATCGCATACCTGAGGCGCTGCTTAAACCTGTGTCTATGCTTTGTC	954	
Qy	978	GG	CTTAAAGTTCGGAAGAACGTCGGAACATTATGAGAGATATCGGCTGCCCTCTCAC	1037	
Db	955	AG	CTGAAGTTTCGAAGAACTTCTGGAACCTTGTGAAGGACATGGTTCCTCCCTTAC	1014	
Qy	1038	CT	GGGAGTCTCAAGTCAATGGAAGTCTCTGAGACAGTTCCAAAGACTGTCTGCTGCTCC	1097	
Db	1015	CT	TGGGCTCTACATCAATGGGAATCTTCTGAGGACAAATCCAAAGACTTTTCTGCTCC	1074	
Qy	1098	CACA	TGTAGAGACCAACAGTATGTTCCAAATGTATAGGCGTTGCCACTCTAGAGAAG	1157	
Db	1075	CACA	ATGTGAGGCCACCAAGATGTTCCA---GTTATAGGCTTGCAGAGGTTTCGAGAA	1131	
Qy	1158	TTA	ATAACAAATCTTAGGAGCATGGCTGATCATTTGGAATGCAACAAGAAAGCTTTGTC	1217	
Db	1132	CT	GCTCTGGAATTTGCAAGTCATGGCTGTGGCTCTTGAATGTGGTGAAGCGCTTTGT	1191	
Qy	1218	TT	ATAGCATGTGAGTATCATGGAGAAG	1245	
Db	1192	TT	ATGCTTGCATTTCTCATGAGAG	1219	

Db 254 CCTTCTGGGCTATGAGGATCCATGATGGATCTTTGGCCAGGTCATGTGCAAGCCC 313
Qy 418 TTCAGGCGATGATACAAATGAATTTACGTGTCATGCTCCTCCTGATCAG 477
Db 314 TACTGGGCGTCTACACTATTAACCTTACACATCCATGCTCCTCCTGATCAGT 373
Qy 478 TGGATCGTTTCATGATGAGTGTGTCAGGCTACCAAGGCTTCAACCGGCGCTAAGTGA 537
Db 374 TGGATCGTTTCATGATGAGTGTGTCAGGCTACCAAGGCTTCAACCGGCGCTAAGTGA 433
Qy 538 AGATCTGGGCGAAGTCAATTTGCTTGTCTTATTTGGGTGGTCTCCCTGTTTCTTTTC 597
Db 434 TGACTTGGGCGAAGTCAATTTGCTTGTCTTATTTGGGTGGTCTCCCTGTTTCTTTTC 493
Qy 598 CACAGATCATCTATGGCAGTGTCAAGATTTAGCAGCTTATCTGTCATGACCAAGTG 657
Db 494 CCCAAATATCTATGGCAATGCTTTAATCTGGCAAGCTCATATGTTGTTTATCATGACA 553
Qy 658 AGAGATATCCACTATGTTCTTCTTATACAGATGACTCTGGGTTCTTCTGCTGCAATTC 717
Db 554 AGGATTTCCACTGTGTTCTTCCACCAGATGACACTGGGTTCTTCTTGGCACTGC 613
Qy 718 TCATATGATCTGTGCTACTCAGGCAATTTCAAGACTTGTCTGTCATGCTGCAACTTC 777
Db 614 TCGCATGATTTGCTGCTATTTCACTATCAATCAAAACACTGCTTCATGCTGGAGGCTTC 673
Qy 778 AGAGACAAATCTTAAAGATCATCTTCTTGTAGTGGTGTCTTCTGCTGACCCAGA 837
Db 674 AGAGACAGATCTCTAAGATCATCTTCTTGTATGGTGTGTTCTGCTGACCCAGA 733
Qy 838 CACCTTTCAACTTGGCATGTTAATCCAAAGTACAAAGTGGAGTACTATACCAATACCA 897
Db 734 CACCTTTCAACTTGGCATGTTAATCCAAAGTACAAAGTGGAGTACTATACCAATACCA 793
Qy 898 GCTTTAGTATGCCATGCTAGTACAGAGCTATAGCATCTTTCGGGCTTGCCTTAACC 957
Db 794 GCTTTACTACCATATAGTACAGAGGCTATCGCATCTTTCGGGCTTGCCTTAACC 853
Qy 958 CTGTACTTTATGCTTTGTTGGCTTAAAGTTTCGGAAGAACGTTCTGGAATCTTATGAAG 1017
Db 854 CTGTGCTATGCTTTGTCAGCTGAAGTTTCGGAAGAACGTTCTGGAATCTTATGAAG 913
Qy 1018 ATATCGGTGCTCTCTCAGCTGGAGTCTCAAGTCAATGAGTCTTCTGAGGAGT 1077
Db 914 ACATTTGGCTGTCTCCCTTACCTTGGGCTTCAATCAATGAAATCTTCTGAGGACAT 973
Qy 1078 CCAAGACTTGTCTGCTCCCAATGTAGAGACCAAGTATGTTTCCAAATTTAG 1133
Db 974 CCAAGACTTGTCTGCTCCCAATGTAGAGACCAAGTATGTTTCCAAATTTAG 1029

RESULT 7

US-09-852-156-3

; Sequence 3, Application US/09852156

; Patent No. US20020076694A1

; GENERAL INFORMATION:

; APPLICANT: Littman, Dan R.

; Deng, Hongkui

; Unutmaz, Derya

; Ramani, Vineet N.K.

; TITLE OF INVENTION: NOVEL ALTERNATIVE G-COUPLED RECEPTORS

; ASSOCIATED WITH RETROVIRAL ENTRY INTO CELLS, METHODS OF

; IDENTIFYING THE SAME, AND DIAGNOSTIC AND THERAPEUTIC USES

; THEREOF

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: David A. Jackson, Esq.

; STREET: 411 Hackensack Ave, Continental Plaza, 4th

; Floor

; CITY: Hackensack

; STATE: New Jersey

; COUNTRY: USA

ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/852,156
FILING DATE: 09-May-2001
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 1049-1-009NCIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ORGANISM: African Green Monkey
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-852-156-3

Query Match 54.8%; Score 685; DB 9; Length 1037;
Best Local Similarity 81.1%; Pred. No. 5e-218;
Matches 809; Conservative 0; Mismatches 185; Indels 3; Gaps 1;
Qy 140 CTTCAACAATTCAGATGATAACAGCCAGGAGGAAACAAC--GCTTCTTAAAGTTCAAGGA 196
Db 33 CTTCAACAGTTTCAATGACAGCAGCAGGAGGAGCATCAGGACTTCTGCAAGTTCAAGAA 92
Qy 197 GGTCTTTTGGCTGTGTACCTGGTGTGTGTCTTTGGAGTGTAGAACTCCCT 256
Db 93 GGTCTTTTGGCTGTGTACCTGGTGTGTGTCTTTGGAGTGTAGAACTCCCT 152
Qy 257 GGTCTGTATTATATACATTTTCTACCAAGCTCAGGACTCTGACAGATGTGTCTGCT 316
Db 153 GGTGTGTGTATATCAATCTTCTACCAAGCTCAGGACTCTGACGAGCTGTCTCTGCT 212
Qy 317 GAATTTGCCCTGGTGTACCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 376
Db 213 GAACTTACCTGGTGTACCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 272
Qy 377 CACCTATCAGTGGGTCTTTGGCAGCATCTGTGCAAACTCTTCGAGGCTATATCAAT 436
Db 273 CATCATGAATGGATCTTTGGCCAGGTCATGTGCAAGCCCTACTGGGTATCTACATAT 332
Qy 437 GAATTTTACGTGTCCATGCTCACTCTCACTGTGATCATCAGTGTGATCGTTTCATGTAGT 496
Db 333 TAACTTTTACATCTATGCTCATCTCACCTGATCATCTGATGATCGTTTTCATGTAGT 392
Qy 497 GGTCCAGGCTACCAAGGCTTCAACCGGCGAGCTAAGTGGAGATCTGGGCGCAAGTCAT 556
Db 393 GGTTAAGGCGCAAGGCTTAAACCAAGCCAGCAAGATGATCTGGGCGCAAGTCAT 452
Qy 557 TTGCTTGTCTTATTTGGGTGGTCTCCCTGTTGGTTTCTTTGGCCACAGATCATCTATGGCCA 616
Db 453 CTGCTTGTCTATCTGGGTGATATCCCTGCTGTTCTTGGCCAAATATCTATGGCAA 512
Qy 617 TGTTCAGATATTGCAAGCTTATCTGTCAGTACCAAGTGTGAGGATATCCATGTGT 676
Db 513 TGTCTTAACTTGGCAAGCTCATATGTTGTTATCATGATGAGGATTTCCACTGTGT 572
Qy 677 TCTTGTATACAGATGACTCTGGGTTCTTCTGCTGCAATGTCTCACTATGATTTCTGTCTA 736
Db 573 TCTTCCACCCAGATGACACTGGGTTCTTCTTGGCACTGTGCGCATGATTTGTCTGTCTA 632

Qy	737	CTCAGGCAATTATCAAGACCTTGCCTTCAATGCTCGAAACATTCAGAGAAGCAAAATCTCTATAA	796
Db	633	TTCAATCATATAAACAACACTGCTTTCATGCTGGAGGCTTCAGAGAAGCAAGATCTCTATAA	692
Qy	797	GATCATCTTCCTTGTAGTGGGTGTGTTCCCTGTGACCCAGACACCCCTTCAACCTTGGCAT	856
Db	693	GATCATCTTCCTTGTGTGTGCTGTGTTCTGCTGACCCAGACACCCCTTCAACCTCGTGA	752
Qy	857	GTTTAATCCAAAGTCAAGCTCGGAGTACTATACCATAAACGAGCTTTAAGTAGTGCATCGT	916
Db	753	GCTCATCCGCAGACACACTCGGAGTACTATGCCATGACCCAGCTTTCACACACCATCAT	812
Qy	917	AGTGACAGAGGCTATAGCATACTTTCGGGCTTGCCTTAACCCCTGTACTTATGCTTTGT	976
Db	813	AGTGACAGAGGCCATCGCATACCTGAGGGCCCTGCCTTAAACCCCTGTACTATGCTTTGT	872
Qy	977	TGGCTTTAAAGTTCGGAGAAGACGTTGTGAAACTTATGAAGGATATCGGCTGCCTCTCTCA	1036
Db	873	CAGCCTGAAGTTTCGAAAGAACTTCGAAACCTTGTGAAGGACATATGGCTGTCTCCCTTA	932
Qy	1037	CCTCGGAGTCTCAAGTCAATCGAAGTCTTCTTGAGGACAGTTCCTCAAGACTTGTCTCTGCGCTC	1096
Db	933	CTTTGGGGTCTCAATCAATGGAATCTTCTTGAGGACAAATTCCAAGACTTTTCTTGCGCTC	992
Qy	1097	CCACAAATGTAGAGACCACGATATGTTCCAAATTGTAG	1133
Db	993	CCACAAATGTGAGGCCACACGACATGTTCCAGTTATAG	1029

RESULT 8
US-10-083-168-3
? Sequence 3, Application US/10083168
? Publication No. US20030023069A1
? GENERAL INFORMATION:
? APPLICANT: Liaw, Chen W.
? APPLICANT: Chalmers, Derek T.
? APPLICANT: Behan, Dominic P.
? APPLICANT: Maciejewski-Lenior, Dominique
? APPLICANT: Leonard, James N.
? APPLICANT: Ortuno, Daniel
? APPLICANT: Lin, I-Lin
? TITLE OF INVENTION: Endogenous And No. US20030023069A1-Endogenous, Constitutively Act
? TITLE OF INVENTION: Receptors
? FILE REFERENCE: AREN-0320
? CURRENT APPLICATION NUMBER: US/10/083,168
? CURRENT FILING DATE: 2002-02-26
? NUMBER OF SEQ ID NOS: 102
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 3
? LENGTH: 1029
? TYPE: DNA
? ORGANISM: Homo sapiens
? US-10-083-168-3

	Query Match	54.4%;	Score 680.8;	DB 15;	Length 1029;
	Best Local Similarity	81.3%;	Pred. No. 1.3e-216;		
	Matches 787;	Conservative 0;	Mismatches 181;	Indels 0;	Gaps 0
Qy	166	AGGAGAACAAACGGTTCCTCAAGTTCAAGGAGGTCTTTTGGCCCTGTGTGTACCTGGTAG	225		
Db	62	AGGAGCATCAAGCCCTTCTCGAGTTTCAGCAAGGCTTTCTGCCCTGCATGTACCTGGTGG	121		
Qy	226	TGTTTGCTCTTTGGACTGCTAGAAACTCCCTGGTTCGTATATATACATTTTCTACACAGA	285		
Db	122	TGTTTGCTGTGGTCTGSGGGAACTCTCTGGTCTGGTCATATCOATCTCTACCATTA	181		
Qy	286	AGCTGAGGAACTCTGACAGATGTGTTTCTGCTGAACTGCCCTGGCTGACCTGGTGGTTG	345		
Db	182	AGTTGCAGAGCCTACGGGATGTGTTCTGGTGAACCTACCCCTGGCTGACCTGGTGGTTG	241		
Qy	346	TCTGTACTCTCCCTTTTGGGCCTATGCAGGACCTATGAGTGGGTCTTTGCGACAGTCA	405		

Db	242	CTTGCACTCTGCCCTTCTGGCCCTATGACAGCATCAATGAATGGGTGTTTGGCCACAGTCA	301
Qy	406	TGTGCAAAACTCTTCGAGGCGATGATACAAATGAACCTTTACGTGTCCATGCTCACTCTCA	465
Db	302	TGTGCAAAAGCCTACTGGGCATCTACACTATTAACCTTCTACAGTCCATGCTCATCTCTCA	361
Qy	466	CCTGTCATCAAGTGGATCGTTTCATTGTAGTGGTCCAGGCTACCAAGGCCTTCAACCGGC	525
Db	362	CCTGTCATCACTGTGGATCGTTTCATTGTAGTGGTAAAGGCCACCAAGGCCTACCAACGAGC	421
Qy	526	AGGCTAAGTGGAAAGATCTGGGGCCAACTCATTTGCTTGCTCATTTGGGTGGTCTCCCTGCT	585
Db	422	AAGCCAGAGGATGACCTGGGGCAAGTCAACAGCTTGCTCACTGGGTGATATCCCTGC	481
Qy	586	TGTTTTCTTTGCCACAGATCATCTATGGCCATGTTCAAGATATTTGACAAAGCTTATCTGTCT	645
Db	482	TGTTTTCTTTGCCCAAATTTATCTATGGCAATGCTTTAATCTCGAAGCTCATATATGT	541
Qy	646	AGTACCACAGTGGAGAGATATCCACTATGCTCTTGTTATACAGATGACACTCTGGGGTCT	705
Db	542	GTACCATGACGAGGCAATTTCCACTGTGGTCTTGCCACCCAGATGACACTGGGGTCT	601
Qy	706	TCCTGCCATTGCTCACATGATGTTCTGTCTACTCAGGCATTATCAAGACCTTGGCTTCATG	765
Db	602	TCITGCCACTGCTCACCATGATTTGCTGTCTATTTCAGTCATAATCAAAACACTGCTTCATG	661
Qy	766	CTCGAAACTTCCAGAGCAAAATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTCTCC	825
Db	662	CTGAGGCTTCCAGAGCACAGATCTCTAAGATCATCTTCTCTGGTGTATGGCTGTGTCTCC	721
Qy	826	TGCTGACCCAGACACCTTCAACCTTGCCCATGTTAATCCAAAGTACAAAGCTGGGAGTACT	885
Db	722	TGCTGACCCAGATGCCCTTCAACCTCATGAAGTTTCATCCGACGACACACTGGGAATACT	781
Qy	886	ATACCATAAACAGCTTTAAGATATGCCATCGTAGTGAACAGGGCTATAGCATACTTTCGGG	945
Db	782	ATGCCATGACCAAGCTTTCATACACCATCATGGTGAACAGAGGCCATCGCATACCTGAGGG	841
Qy	946	CTTGCCCTTAACCTGTACTTTATGCTCTTGTGGCTTAAAGTTCGGAAGAACAGCTCTGGA	1005
Db	842	CCTGCCCTAAACCTGTGCTCTATGCTCTTGTCAAGCTTGAAGTTTTCGAAAGAACTCTGGA	901
Qy	1006	AACCTATGAAGGATATCGGTGCTCTCTCACCTGGGAGTCTCAAGTCAATGGAAGTCTT	1065
Db	902	AACCTTGTGAAGGACATTTGGTTGGCTTCCCTTACCTTGGGGTCTCACATCATTGGAAATCTT	961
Qy	1066	CTGAGGACAGTTCCAAGACTTGTTCGTCCCTCCCACAATGTAGAGACCAACAGATATGTTC	1125
Db	962	CTGAGGACAATTCGAAGACTTTTCTGCCTCCCAACAATGTGGAGGCCACCAAGCATGTTC	1021
Qy	1126	AATTGTAG	1133
Db	1022	AGTTATAG	1029

```

RESULT 9
US-09-940-063-1
; Sequence 1, Application US/09940063
; Patent No. US20020090657A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; APPLICANT: Murphy, Kristine E.
; APPLICANT: Wilbanks, Alyson M.
; APPLICANT: Wu, Lijun
; TITLE OF INVENTION: US20020090657A1el Antibodies and Ligands for "Bonzo"
; TITLE OF INVENTION: Chemokine Receptor
; FILE REFERENCE: 1855.1070-060
; CURRENT APPLICATION NUMBER: US/09/940,063
; CURRENT FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 09/449,437
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 1
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-940-063-1

```
Query Match      54.3%; Score 679.2; DB 9; Length 1029;  
Best Local Similarity 81.2%; Pred. No. 4.4e-216;  
Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;  
  
QY 166 AGGAGAACAAAGCGTCTCTAAAGTTCAAGGAGGTCCTTTTGGCCCTGTGTCACCTGCTAG 225  
DB 62 AGGAGCATCAAGACTCTCTCGAGTTCAAGCAAGGTCCTTCTGCCCTGATGCTACCTGGTGG 121  
  
QY 226 TGTGTGCTTTGGACTCTAGGAACCTCCTGTTCTGATTATATACATTTCTACACAGA 285  
DB 122 TGTGTGCTGTGGTCTCTGGTGGGAACCTCTCTGGTCTGTCATATCCATCTTACCATTA 181  
  
QY 286 AGCTGAGGACTCTGACAGATGTGTTTCTGCTGAACCTTGGCCCTGGGTGACCTGGTGTG 345  
DB 182 AGTTGCAGAGCCTGACGGATGTCTCTGGTGAACCTTACCCTGGCTGACCTGTGTGTTG 241  
  
QY 346 TCTGTACTCTGCCCTTTTGGSCCTATCGACGACCTTAAGAGTGGTCTTTTGGCACAGTCA 405  
DB 242 TCTGCACCTCTGCCCTTTTGGSCCTATCGACGACCTTCAATGAATGGGTGTTTGGGCCAGTCA 301  
  
QY 406 TGTCAAAACCTCTTCGAGGATGTATACAATGAACCTTCTACGTGTCATCTCACTCTCA 465  
DB 302 TGTGCAAGAGCCTACTGGGATCTACACTATTAACCTTCTACAGCTCCATGCTCATCTCA 361  
  
QY 466 CCTGCATCAGAGTGGATCGTTTCAATGTAGTGTCCAGGTACCAAGGCTTCAACCGGC 525  
DB 362 CCTGCATCAGTGGATCGTTTCAATGTAGTGTCCAGGTACCAAGGCTTCAACCGGC 421  
  
QY 526 AGGTAAGTGGAGATCTGGGCGCAAGTCATTTGCTGCTCATTTGGGTGGTCTCCCTGT 585  
DB 422 AAGCCAAAGAGGATCACTTGGGCGCAAGTCATTTGCTGCTCATTTGGGTGGTATCCCTGC 481  
  
QY 586 TGTGTTCTTTGGCCAGATGTATCAACTATGTTCTTGTATATACAGATGACTCTGGGTCT 705  
DB 482 TGTGTTCTTTGGCCAGATGTATCAACTATGTTCTTGTATATACAGATGACTCTGGGTCT 601  
  
QY 706 TCTGCAATGTCTCAATGATCTGCTGCTACTCAGGCAATATCAAGACCTTCTGCTCATG 765  
DB 602 TCTTGCACACTGCTCAACATGATGTCTGCTATTTACGTCATAATCAAACTGCTTCATG 661  
  
QY 766 CTGGAATCTTCCAGAGCACAATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTCC 825  
DB 662 CTGAGGCTTCCAGAGCACAAGATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTCC 721  
  
QY 826 TGTGACCCAGACACCTTCAACCTTGGCATGTAAATCCAAAGTACAAAGCTGGGAGTACT 885  
DB 722 TGTGACCCAGATGCCCCCTTCAACCTATGAAGTTTATCCACAGCACACACTGGGAATCT 781  
  
QY 886 ATACCATAAACAGCTTTAAGTATGCCATCGTAGTGAAGAGGCTATAGCATATCTTCGGG 945  
DB 782 ATGCCATGACAGCTTTTCACTACACATCATGTGTGACAGAGGCTATCGCATCTCAGGG 841  
  
QY 946 CTTCGCTTAACCTGTACTTTATGCTTTGCTTGGCTTAAGTTCCGAGAGACGCTGGA 1005  
DB 842 CTTCGCTTAACCTGTCTCTATGCTTTTGTCAAGCTTGAAGTTTCGAAAGAACTTCTGGA 901  
  
QY 1006 AACTTATGAAGGATATCGGCTGCTCTCTCACCTGGAGTCTCAAGTCAATGGAAGTCTT 1065  
DB 902 AACTTGTGAAGGACATTTGCTCTCTTACCTTGGGTCTCACTCAATGGAATCTT 961  
  
QY 1066 CTGAGGACAGTTCGAAGACTTTTCTGCTCCCAATGTAGAGACACCAAGTATGTTCC 1125  
DB 962 CTGAGGACAAATCCGAAGCTTTTCTGCTCCCAATGTGAGGAGGCCACCAAGATGTTCC 1021
```

QY 1126 AATTGTAG 1133
DB 1022 AGTTATAG 1029

RESULT 10
US-10-174-293-1
; Sequence 1, Application US/10174293
; Publication No. US20030165995A1
; GENERAL INFORMATION:
; APPLICANT: Briskin, Michael J.
; APPLICANT: Murphy, Kristine E.
; APPLICANT: Wilbanks, Alyson M.
; APPLICANT: Wu, Lijun
; TITLE OF INVENTION: No. US20030165995A1el Antibodies and ligands for "Bonzo"
; FILE REFERENCE: 1855.1070-001
; CURRENT APPLICATION NUMBER: US/10/174,293
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US/09/722,064
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 09/449,437
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-174-293-1

```
Query Match      54.3%; Score 679.2; DB 15; Length 1029;  
Best Local Similarity 81.2%; Pred. No. 4.4e-216;  
Matches 786; Conservative 0; Mismatches 182; Indels 0; Gaps 0;  
  
QY 166 AGGAGAACAAAGCGTCTCTAAAGTTCAAGGAGGTCCTTTTGGCCCTGTGTCACCTGCTAG 225  
DB 62 AGGAGCATCAAGACTCTCTCGAGTTCAAGCAAGGTCCTTCTGCCCTGATGCTACCTGGTGG 121  
  
QY 226 TGTGTGCTTTGGACTCTAGGAACCTCCTGTTCTGATTATATACATTTCTACACAGA 285  
DB 122 TGTGTGCTGTGGTCTCTGGTGGGAACCTCTCTGGTCTGTCATATCCATCTTACCATTA 181  
  
QY 286 AGCTGAGGACTCTGACAGATGTGTTTCTGCTGAACCTTGGCCCTGGCTGACCTGGTGTG 345  
DB 182 AGTTGCAGAGCCTGACGGATGTGTTTCTGCTGAACCTTACCCCTGGCTGACCTGGTGTG 241  
  
QY 346 TCTGTACTCTGCCCTTTTGGSCCTATCGACGACCTTATGAGTGGGTCTTTGGCACAGTCA 405  
DB 242 TCTGCACCTCTGCCCTTTTGGSCCTATCGACGACCTTATGAGTGGGTGTTTGGCCAGTCA 301  
  
QY 406 TGTCAAAACCTCTTTCGAGGCAATGTATACAATGAACCTTCTACGTGTCATGCTCACTCTCA 465  
DB 302 TGTGCAAGAGCCTACTGGGATCTACACTATTAACCTTCTACAGCTCCATGCTCATCTCA 361  
  
QY 466 CCTGCATCAGAGTGGATCGTTTCAATGTAGTGTCCAGGTACCAAGGCTTCAACCGGC 525  
DB 362 CCTGCATCAGTGGATCGTTTCAATGTAGTGTTCAGGCTTCAAGATATGCAAGCTTATCTGC 645  
  
QY 526 AGGCTAAAGTGAAGATCTGGGCGCAAGTCATTTGCTGCTCATTTGGGTGGTGTCTCCCTGT 585  
DB 422 AAGCCAAAGAGGATCACTTGGGCGCAAGTCATTTGCTGCTCATTTGGGTGGTATTCCTGC 481  
  
QY 586 TGTGTTCTTTGGCCAGATGTATCAACTATGTTCTTGTATATACAGATGACTCTGGGTCT 705  
DB 482 TGTGTTCTTTGGCCAGATGTATCAACTATGTTCTTGTATATACAGATGACTCTGGGTCT 601  
  
QY 646 AGTACCAAGTGAAGAGATATCCACTATGTTCTTGTATATACAGATGACTCTGGGTCT 705  
DB 542 GTTACCATGACAGGCAATTTCCACTGTGGTCTTGGCCACCCAGATGACACTGGGTCT 601  
  
QY 706 TCTGCCCAATGCTCACTATGATTTCTGTCTACTCAGGCAATATCAAGACCTTCTGCTCATG 765
```

```
Db 602 TCTTGCACCTGCTCACCATGATGTCGTCTATTTCAGTCATTAATCAAAACATGCTTCAATG 661
Qy 766 CTCGAAATTCAGAGACAAATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTC 825
Db 662 CTGGAGGCTTCAGAGACACAGATCTCTAAAGATCATCTTCTTGTAGTGGCTGTGTTC 721
Qy 826 TGCTGACCCAGACACCCCTCAACCTTGCATGTTAATCAAGTACAGCTGGAGTACT 885
Db 722 TGCTGACCCAGATGCCCTTCAACCTCATGAAGTTTCATCCGACGACACACTGGGAATACT 781
Qy 886 ATACCATAACCAAGCTTTAAGTAGTGCATGCTAGTACAGAGGCTATAGCATACTTTCGGG 945
Db 782 ATGCCATGACCAAGCTTTCACATACACCATCATGCTGACAGAGGCCATGACATACCTTGAGG 841
Qy 946 CTGTCCTTAACCCCTGACTTATGCTTATGCTTGTGCTTAAAGTTCCGGAAGACGCTGCGA 1005
Db 842 CTGTCCTTAACCCCTGCTCTATGCTTGTGCTGCTGAAGTTTCGGAAGAACTTCTGGA 901
Qy 1006 AACTTATGAAGGATATCGGCTGCCCTCTCTCACTCGGAGTCTCAAGTCAATGGAGTCTT 1065
Db 902 AACTTGTGAGACATGCTGCTGCCCTTACCTTGGGGTCTCACATCAATGGAACTT 961
Qy 1066 CTGAGACAGTTCAGAGCTTCTTCTGCTCCCAAAATGTAGAGACCAAGTATGTTC 1125
Db 962 CTGAGGACAAATCCAGAGCTTCTTCTGCTCCCAAAATGTAGAGACCAAGTATGTTC 1021
Qy 1126 AATTGTAG 1133
Db 1022 AGTTATAG 1029

RESULT 11
US-10-353-690-5
; Sequence 5, Application US/10353690
; Publication NO. US20030215840A1
; GENERAL INFORMATION:
; APPLICANT: Logan, Thomas Joseph
; APPLICANT: Chun, Miyoung
; APPLICANT: Galvin, Katherine M.
; APPLICANT: Healy, Aileen
; APPLICANT: Acton, Susan L.
; APPLICANT: Donoghue, Mary
; APPLICANT: Stagliano, Nancy
; APPLICANT: Perodin, Jacqueline
; APPLICANT: Rodriguez-Way, Amelie
; TITLE OF INVENTION: Methods and compositions for treating
; TITLE OF INVENTION: cardiovascular disease using 1682, 6169, 6193, 7771, 14395,
; TITLE OF INVENTION: 28002, 33215, 43726, 59292, 26156, 32427, 2402, 7747, 1720,
; TITLE OF INVENTION: 9151, 60491, 1371, 7077, 33207, 1419, 18036, 16105, 38650,
; TITLE OF INVENTION: 14245, 58848, 1870, 25856, 32394, 3484, 345, 5252, 9135,
; TITLE OF INVENTION: 10532, 18610, 8165, 2448, 2445, 46424, 84237, 8912, 2868,
; TITLE OF INVENTION: 283, 2554, 9464, 17799, 26686, 43848, 32135, 12208, 2914,
; TITLE OF INVENTION: 51130, 19489, 21833, 2917, 59590, 15992, 2094, 2252, 3474,
; FILE REFERENCE: MPI02-018P1R0M1M
; CURRENT APPLICATION NUMBER: US/10/353,690
; CURRENT FILING DATE: 2003-01-29
; PRIOR APPLICATION NUMBER: 60/353,224
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/364,529
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: 60/373,861
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/376,287
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: 60/388,080
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: 60/390,971
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/394,130
; PRIOR FILING DATE: 2002-07-03
; PRIOR APPLICATION NUMBER: 60/394,797
```

```
; PRIOR FILING DATE: 2002-07-10
; PRIOR APPLICATION NUMBER: 60/404,904
; PRIOR FILING DATE: 2002-08-21
; PRIOR APPLICATION NUMBER: 60/405,450
; PRIOR FILING DATE: 2002-08-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Homo Sapiens
US-10-353-690-5

Query Match 54.3% Score 679.2; DB 16; Length 1029;
Best Local Similarity 81.2%; Pred. No. 4.4e-216; Mismatches 182; Indels 0; Gaps 0;
Matches 786; Conservative 0;

Qy 166 AGGAGAACAAACGCTTCTAAAGTTCAAGAGAGTCTTTTGGCCCTGTGTGTACCTGGTAG 225
Db 62 AGGAGCATCAAGACTTCTGCAGTTTCAGCAAGGTCTTTCTGCCCTGCATGTACTGGTGG 121
Qy 226 TGTGTGCTTTGGAGCTGCTAGGAACTCCCTGCTGTTCTGATTATATACATTTTCTACAGA 285
Db 122 TGTGTGCTGCTGTGGTGGGAACTCTCTGTGTGTGTCATCTCATCTTCTTACCATA 181
Qy 286 AGCTGAGGACTCTGACAGATGTGTTCTGCTGAACTTGCCTTGGCTGGCTGACCTGGTGTG 345
Db 182 AGTTGCAGAGCTTGCAGGATGCTTCTGCTGTAACCTACCCCTGGCTGACCTGGTGTG 241
Qy 346 TGTGTACTTGCCTTTTGGGCTTATGAGGACCTATGAGTGGGCTTTGGGACAGTCA 405
Db 242 TCTGCACTTGCCTTCTGGGCTTATGAGGACCTATGAGTGGGCTTTGGGACAGTCA 301
Qy 406 TGTGCAAACTCTTCGAGGACATGTATACATGAATTTCTACGTGTCATCTCACTCTCA 465
Db 302 TGTGCAAGAGCTTACTGGGATCTACACATTAATTTCTACAGCTGCTCATCTCA 361
Qy 466 CTGTCATCAGTGGATGCTTTTCTTGTAGTGTGTCAGGCTACCAAGGCTTCAACCGGC 525
Db 362 CTGTCATCAGTGGATGCTTTTCTTGTAGTGTGTAAGGCCACCAAGGCTTCAACCGGC 421
Qy 526 AGCTAGTGGAGAGTCTGGGCTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 585
Db 422 AGCCAAAGAGATGACCTGGGCAAGTCAACAGCTTGTCTGCTGCTGCTGCTGCTGCTG 481
Qy 586 TGTGTTCTTTGCCACAGATCATCTATGGCCATGTTCAAGATATTGACAAAGCTTATCTGTC 645
Db 482 TGTGTTCTTTGCCCAAAATTTATCTATGGCAATGCTTTTAACTCTCGACAAGCTCATATGTG 541
Qy 646 AGTACCAAGTGGAGAGATATCCACTATGCTTCTTGTATACAGATGACTCTGGGTTCT 705
Db 542 GTTACCAATGACGAGGCAATTTCCACTGTGTTCTTGCACCCAGATGACACTGGGTTCT 601
Qy 706 TCTGTCATTTGCTCAGTATGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 765
Db 602 TCTTGCACATGCTCACCATGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 661
Qy 766 CTGAAACTTCCAGAGACAAATCTCTAAAGATGCTTCTTCTGCTGCTGCTGCTGCTGCT 825
Db 662 CTGGAGGCTTCCAGAGACAGATCTCTAAAGATCATCTTCTTCTGCTGCTGCTGCTGCTG 721
Qy 826 TGCTGACCCAGACACCTTCAACCTTGGCCATGTTAATCCAAAGTACAAAGTGGGAGTACT 885
Db 722 TGCTGACCCAGATGCCCTTCAACCTCATGAAGTTTCATCCGACGACACACTGGGAATACT 781
Qy 886 ATACCATAACCAAGCTTTAAGTAGTGCATGCTAGTACAGAGGCTATAGCATACTTTCGGG 945
Db 782 ATGCCATGACCAAGCTTTCACATACACCATCATGCTGACAGAGGCCATGCTGAGGG 841
Qy 946 CTGTCCTTAACCCCTGACTTATGCTTATGCTTGTGCTTAAAGTTCCGGAAGACGCTGCGA 1005
Db 842 CTGTCCTTAACCCCTGCTCTATGCTTGTGCTGCTGAAGTTTCGGAAGAACTTCTGGA 901
```

Qy	1006	AAC	TATGAAGGATATCGG	TGCTCTCTCACTGGGAGTCTCAAGTCAATGGAAGTCTT	1065
Db	902	AAC	TGTGAAGGACATTTGG	TTCCTCCTTACCTTGGGGTCTCACATCAATGGAAATCTT	961
Qy	1066	CTGAGGACAGTTTCA	AGACATTTGTTCTGCTCCACCAATGTAGAGACACCAGATGTTC	1125	
Db	962	CTGAGGACAAATTT	CCAAGACTTTTCTGCTCCCAATGTGGAGGCCACCAGCATGTTC	1021	
Qy	1126	AAT	TGTAG	1133	
Db	1022	AGT	TATAG	1029	

RESULT 12

US-10-083-168-64

; Sequence 64, Application US/10083168

; Publication No. US20030023069A1

; GENERAL INFORMATION:

; APPLICANT: Liaw, Chen W.

; APPLICANT: Chalmers, Derek T.

; APPLICANT: Benan, Dominic P.

; APPLICANT: Maciejewski-Lenior, Dominique

; APPLICANT: Leonard, James N.

; APPLICANT: Ortuno, Daniel

; APPLICANT: Lin, I-Lin

; TITLE OF INVENTION: Endogenous And No. US20030023069A1-Endogenous, Constitut

; TITLE OF INVENTION: Receptors

; FILE REFERENCE: AREN-0320

; CURRENT APPLICATION NUMBER: US/10/083,168

; CURRENT FILING DATE: 2002-02-26

; NUMBER OF SEQ ID NOS: 102

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 64

; LENGTH: 1029

; TYPE: DNA

; ORGANISM: Unknown

; FEATURE:

; OTHER INFORMATION: No. US20030023069A1el Sequence

US-10-083-168-64

RESULT 13
US-09-852-156-1
Sequence 1, Application US/09852156
Patent No. US20020076694A1
GENERAL INFORMATION:
APPLICANT: Littman, Dan R.
Deng, Hongkui
Unutmaz, Derya
Ramani, Vineet N.K.
TITLE OF INVENTION: NOVEL ALTERNATIVE G-COUPLED RECEPTORS
ASSOCIATED WITH RETROVIRAL ENTRY INTO CELLS, METHODS OF
IDENTIFYING THE SAME, AND DIAGNOSTIC AND THERAPEUTIC USES
THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
Floor
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/852,156
FILING DATE: 09-May-2001
CLASSIFICATION: <unknown>
ATTORNEY/AGENT INFORMATION:

Db	751	TTCATCCGAGCACACACTGGGAATACTATGTGCATGACCAGCTTTTCACTACACCATCATG	811
Qy	918	GTCACAGAGCTATAGCATACTTTCCGGCTTGCCTTAAACCCCTGTACTTTATTGCTTTGTT	977
Db	811	GTCACAGAGGCCATCGCATACCTTGAGGGCCCTGCTTAACCCCTGTCTATGCTCTTTGTC	870
Qy	978	GGCTTAAAGTTCGGGAAGAACGCTCTGAAAACCTTATGAAGGATATCGGCTGCCTCTCTCAC	1037
Db	871	AGCCTGAAGTTTCGAAAGRACTTCTGAAAACCTTGTGAAGGACATTTGGTTGCCTCCCTTAC	930
Qy	1038	CTCGGAGTCTCAAGTCAATGGAAGTCTTCTGAGGACAGTTTCCAAGACTTGTCTGCGCTCC	1097
Db	931	CTTTGGGGTCTCACATCAATGGAATCTTCTGAGGACAATTTCAAAGACTTTTTTCTGCTCC	990
Qy	1098	CACAATGTAGAGACCACCACTAGTATGTTCCAAATTGTAG	1133
Db	991	CACAATGTGAGGCCACCAGCATGTTCCACTTATAG	1026
RESULT 14			
US-10-251-385-23			
; Sequence 23, Application US/10251385			
; Publication No. US20030105292A1			
; GENERAL INFORMATION:			
; APPLICANT: Behan, Dominic P.			
; APPLICANT: Chalmers, Derek T.			
; APPLICANT: Liaw, Chen W.			
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated			
; TITLE OF INVENTION: Protein-Coupled			
; TITLE OF INVENTION: Receptors			
; FILE REFERENCE: AREN-0040			
; CURRENT APPLICATION NUMBER: US/10/251,385			
; CURRENT FILING DATE: 2002-09-20			
; PRIOR APPLICATION NUMBER: US/09/170,496			
; PRIOR FILING DATE: 1998-10-13			
; NUMBER OF SEQ ID NOS: 294			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 23			
; LENGTH: 1074			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-251-385-23			
Query Match 14.7%; Score 183.4; DB 15; Length 1074;			
Best Local Similarity 52.9%; Pred. No. 7.5e-50;			
Matches 490; Conservative 0; Mismatches 401; Indels 36; Gaps 3			
Qy	146	CAATTCCAGTGTAAACGCCAGGAGAACAAACCGCTTCTTAAGTTCAAGGAGGTCTTTTT	205
Db	57	CAACTTCACTGACTTCTACTGTGAGAAAAACAAATCAGGCAGTTTGGAGCCATTTCCT	116
Qy	206	GCCCTGTGTACTCGTAGT	265
Db	117	CCCACCCCTTTACTCGGCTCGTGTTTCACTCGTGGGTGCCCTTGGGCAACAGTCTTGTTATCCT	176
Qy	266	TATATACATTTTCTACCAAGAGCTTGAGGACTCTGACAGATGTGTTTCTGTGAACTTGCC	325
Db	177	TGTTCTACTGTACTGCAAGAGTGAAGACCATCAGGCATGTTTCCITTTTGAATTTGGC	236
Qy	326	CCTGGCTGACCTGGT	385
Db	237	AATTGCTGACCTCTCTTCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	296
Qy	386	GTGGGTCTTTGGCAGCTCATGTGCAAAACCTTTGAGGCATGTATACAAATGAATCTCTA	445
Db	297	GTGAAAGTTCCAGACCTTCACTGCAAGGTGGTCAACAGCATGTACAGAGATGAATCTCTA	356
Qy	446	CGTGTCCATGCTACTCTCACTCGATCAGTGGATCGTTTCAATGTAGTGTGTGTGTGTGTGT	505
Db	357	CAGCTGT	416
Qy	506	TACCAAGCCTTCAACGGCAGGCTAGTGTGAAGATCTGGGGCCCAAGTCAATTTGCTTGTCT	565

Db 417 CATGAGACACATACTTGGAGGAGAAAAAGGCTTTGTACAGCAAAATGGTTTGCCTTAC 476
Qy 566 CATTGTTGGTGGTCTCCCTGTTGGTTCTTTGCCACAGATCATCTATGGCCATGTTCAAGA 625
Db 477 CATCTGGGTATGGCAGCTGCTCTGTCATCCCAAGAAATCTTATACAGCCAAATCAAGGA 536
Qy 626 T-----ATTGACAAAGCTTATCTGTCACTACACAGTGGAGAGATATCCCATATG-- 674
Db 537 GGAATCCGGCATTTGCTATCTGCACCATGTTTACCTAGCGATGAGAGCACCAACTGAA 596
Qy 675 -----GTTCTGTTATACAGATGACTCTGGGGTTCTTCTGCGATTTGCTCACTATGAT 727
Db 597 GTCAGCTGTCTGACCTTGAAGGTCATCTCTGGGGTTCTTCTTCCCTTCGGTGCATGGC 656
Qy 728 TCTGTGCTACTCAGGCATTTATCAAGACCTTCTGCTCATGCTCGAACTTCCAGAGACCAA 787
Db 657 TTGCTGCTATACCATCATCTACACCTGATACAGCCAGAGTCTTCCAGACCAA 716
Qy 788 ATCTCTAAGATCATCTTCTTCTGTAGTGGTGTGTTCTGCTGACCCAGACACCCCTTCAA 847
Db 717 AGCCCTAAAAGTGACCATCATCTGCTGACCGTCTTTGTCTGTCTCAGTTTCCCTACAA 776
Qy 848 CTTTGCCATGTTAATCCAAAGTCAAGCTGGAGTACTATACCATACAGCTTTTAAGTA 907
Db 777 CTGCATTTTGTGTGACACCATGACGCTATGCCATGTTCAITCTCCAACTGTGCCGT 836
Qy 908 TGCCA-----TCGTAGTACAGAGGCTATAGCATACTTTCGGGGTTG 949
Db 837 TTCCACCAACATGACATCTGCTCCAGGTCAACAGACCATCGCTTCTTCCACAGTTG 896
Qy 950 CTTTAACCTGACTTTATGCTTGTGCTTAAAGTTCGGAGAACGCTCTGGAACCT 1009
Db 897 CTTGAACCTGTTCTCTATGTTTGTGGGTGAGAGATTCGGCGGATCTCGTGA AAC 956
Qy 1010 TATGAAGATATCGGCTGCTCTCTCA 1036
Db 957 CCTGAAGAACTTGGTTGTCATCAGCCA 983

RESULT 15
US-10-225-567A-240
; Sequence 240, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 240
; LENGTH: 2462
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-240

Query Match 14.7%; Score 183.4; DB 15; Length 2462;
Best Local Similarity 52.9%; Pred. No. 1.3e-49;
Matches 490; Conservative 0; Mismatches 401; Indels 36; Gaps 3;

Qy 146 CAATTCAGTATTAACAGCCAGGAGACAAACGTTCTCTAAAGTTCAAGAGGCTCTTTT 205
Db 169 CAACTTCAGTACTTCTACTGTGAGAAAAACAATGTGAGGAGTTTTCGAGCCATTTCT 228
Qy 206 GCCCTGTGTGACTGCTGTGTTGTTGTTGTTGACTCTAGGAAACCTCCCTGGTTCCTGAT 265
Db 229 CCCACCTTGTACTGGCTGTGTTTCATGCTGGGTGCTTGGGCAACAGTCTTGTATCTCT 288

Search completed: September 29, 2004, 03:12:52
Job time : 541 secs

Qy 266 TATATACATTTTCTACCAAGCTGAGGACTCTGACAGATGTTGTTTCTGCTGAACCTTGC 325
Db 289 TGTCTACTGTTACTGCACAAGAGTGAAGCAAGACCGACATGTTTCTTTTGAATTTGGC 348
Qy 326 CTTGGCTGACCTTGTGTTTGTCTGTACTCTGCTTCTGCTTCTTGGGCTATGAGGACCTATGA 385
Db 349 AATTGCTGACCTCTCTTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 408
Qy 386 GTGGGCTTTTGGACAGTATGTCACAAACTCTTCGAGGCATGTATACAAATGAACCTTTTA 445
Db 409 GTGAAGTTTCCAGACTTCTCATGTCAAGGTGTCAACAGCATGTATCAAGATGAACCTTTTA 468
Qy 446 CQTGTCCATGCTCACTCTCACCTGCATCACAGTGTGATGTTTTCATTGTAGTGGTCCAGGC 505
Db 469 CAGCTGTGTTGCTGATCATGTGCATCAGCGTGCACAGGTACATTGCCATTGCCAGGC 528
Qy 506 TACCAAGGCTTCAACCGGACAGGCTAAGTGAAGATCTGGGGCCAAAGTCATTTCCTTGGCT 565
Db 529 CATGAGACACATACTTGGAGGGAGAAAAGGCTTTGTACAGCAAAATGGTTTGTCTTTAC 588
Qy 566 CATTGSGGTGCTCCCTGTTGTTTCTTGTCCACAGATCATCTATGCTGCTTCAAGA 625
Db 589 CATCTGGGTATTTGGCAGCTGCTCTGTCATCCAGAAATCTTATACAGCCAAATCAAGGA 648
Qy 626 T-----ATTGACAAAGCTTATCTGTCAGTACACAGTGGAGAGATATCCACTATG-- 674
Db 649 GGAATCCGGCATTTGCTATCTGCACCATGTTTACCTAGCGATGAGAGCACCAAACTGAA 708
Qy 675 -----GTTCTGTTATACAGTACTCTGGGGTTCTTCTGCTGCTGCTGCTGCTGCTGCT 727
Db 709 GTCAGTGTCTTGAACCTGAAAGTCATCTGGGGTTCTTCTCTTCCCTTCGTGCTATGGC 768
Qy 728 TCTGTGCTACTCAGGCATTTATCAAGACCTTGTCTTCATGCTCGAAACTTCCAGAGACCAA 787
Db 769 TTGCTGCTATACCATCATCATCACTTCAACCTGATACAGCCCAAGAGTCTTCCAGACCAA 828
Qy 788 ATCTCTAAGATCATCTTCTTGTAGTGGCTGTGTTCTGCTGACCCAGACACCCCTTCAA 847
Db 829 AGCCCTAAAAGTGACCATCACTGCTGACCGTCTTTGTCTGTCTCAGTTTCCCTACAA 888
Qy 848 CTTTGCCATGTTAATCCAAAGTACAAAGCTGGAGTACTATACCATACCAAGCTTAAAGTA 907
Db 889 CTGCATTTTGTGGTGCAGACCATTTGACGCTATGCCATGTTCACTCCAACCTGCGGT 948
Qy 908 TGCCA-----TCGTAGTACAGAGGCTATAGCATACTTTCGGGCTTG 949
Db 949 TTCCACCAACATTTGACATCTGCTTCCAGGTCAACCCAGACCATCGCTTCTTCCACAGTTG 1008
Qy 950 CTTTAACCTGTACTTATGCTTGTGCTTAAAGTTCGGGAGAACGCTCTGGAACCT 1009
Db 1009 CTTGAACCTGTTCTCTATGTTTTTGTGGGTGAGAGATTCGCGGGATCTCTGTGAAAAC 1068
Qy 1010 TATGAAGATATCGGCTGCTCTCTCA 1036
Db 1069 CCTGAAGAACTTGGTTGTCATCAGCCA 1095